DIAGNOSIS OF SHOULDER PAIN IN PRIMARY CARE

Angela Cadogan

A thesis submitted to AUT University in fulfilment of the requirements for the degree of
Doctor of Philosophy (PhD)

2011

School of Rehabilitation and Occupation Studies

Faculty of Health and Environmental Science

TABLE OF CONTENTS

ATTESTATION OF AUTHORSHIP

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent, has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signed: ....

Dated: ....

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND GLOSSARY

Abbreviations

+LR	positive likelihood ratio
-LR	negative likelihood ratio
ACJ	acromioclavicular joint
AOR	adjusted odds ratio
AUC	area under the curve
CAL	coracoacromial ligament
CSI	corticosteroid injection
ERLS	external rotation lag sign
GHJ	glenohumeral joint
HHD	hand-held dynamometer
ICC	intraclass correlation coefficient
LOA	limits of agreement
MRI	magnetic resonance imaging
MRA	magnetic resonance arthrogram
NAR	negative anaesthetic response
NPV	negative predictive value
OST	orthopaedic special tests
PABAK	prevalence and bias adjusted kappa
PAR	positive anaesthetic response
PPV	positive predictive value
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCT	rotator cuff tear
ROM	range of movement
ROC	receiver operator curve

SAB   subacromial bursa

SLAP   superior labrum anterior to posterior

STARD   Standards for Reporting of Diagnostic Accuracy

VAS   visual analogue scale

Glossary of Terms

Adjusted odds ratio   odds of having the condition when considered among a number of specific predictor variables (in regression analyses)

Negative likelihood ratio   likelihood of a negative test in those with the condition compared to those without the condition.

Negative predictive value   probability that someone with a negative test won't have the condition

Odds ratio   measure of effect size describing the strength of association between two binary variables.

Positive likelihood ratio   likelihood of a positive test in those with the condition compared to those without the condition.

Positive predictive value   (also called post-test probability). Probability that someone with a positive test will have the condition

Primary health care   primary contact practitioners including general practitioners and physiotherapists.

Secondary care   orthopaedic specialist, surgical centres.

Sensitivity   probability of test being positive when the condition is present. (100% sensitivity: negative test will rule-out the condition)

Specificity   probability of test being negative when the condition is absent. (100% specificity: positive test will rule-in the condition).
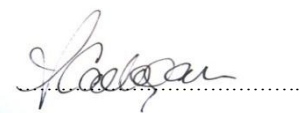
DECLARATION OF AUTHOR CONTRIBUTION

All the co-authors on the chapters/papers indicated in the following table have approved these works for inclusion in Angela Cadogan's doctoral thesis.
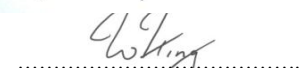
Cadogan, A., Laslett, M., Hing, W., McNair, P., Williams, M.(2011). Reliability of a new hand-held dynamometer in measuring shoulder range of motion and strength. *Manual Therapy*, 16, 97-101.	AC 80%, MW 5%, WH 5%, ML 5%, PM 5%.
Cadogan, A., Laslett, M., Hing, W.A., McNair, P.J., Williams, M. (2011). Interexaminer reliability of orthopaedic special tests used in the assessment of shoulder pain. *Manual Therapy,* 16, 131-135.	AC 80%, MW 5%, WH 5%, ML 5%, PM 5%.
Cadogan, A., Laslett, M., Hing, W., McNair, P., Coates, M.(2011). A prospective study of shoulder pain in primary care: Prevalence of imaged pathology and response to guided diagnostic blocks. *BMC Musculoskeletal Disorders,* 12, 119.	AC 80%, MC, 5%, WH 5%, ML 5%, PM 5%.
Cadogan, A., Laslett, M., Hing, W.A., McNair, P.J., Taylor, S. Clinical predictors of a positive anaesthetic response to guided diagnostic block into the subacromial bursa. (*submitted*)	AC 80%, ST 10%, ML 3%, WH 3%, PM 3%.
Cadogan, A., Laslett, M., Hing, W.A., McNair, P.J., Taylor, S.. Clinical predictors of a positive response to guided diagnostic block into the acromioclavicular joint. (*under review*).	AC 80%, ML 5%, WH 5%, PM 5%, ST 5%.

Signatures:

Angela Cadogan (doctoral candidate)
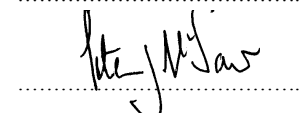
Wayne Hing (primary supervisor)
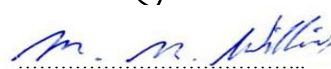
Dr Mark Laslett (secondary supervisor)

Dr Peter McNair (secondary supervisor)

Maynard Williams

Steve Taylor

Dr Mark Coates

RESEARCH OUTPUTS RESULTING FROM THIS THESIS

Peer-Reviewed Publications

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Williams, M. (2011). Interexaminer reliability of orthopaedic special tests used in the assessment of shoulder pain. *Manual Therapy, 16*, 131-135.

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Williams, M. (2011). Reliability of a new hand-held dynamometer in measuring shoulder range of motion and strength. *Manual Therapy, 16*(1), 97-101.

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Coates, M. (2011). A prospective study of shoulder pain in primary care; Prevalence of imaged pathology and response to guided diagnostic blocks. *BMC Musculoskeletal Disorders, 12*, 119.

Manuscripts Submitted

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Taylor, S. (submitted). Clinical predictors of a positive anaesthetic response to guided diagnostic block into the subacromial bursa. *(Journal of Rehabilitation Medicine)*

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Taylor, S. (under review). Clinical diagnosis of a positive response to guided acromioclavicular joint diagnostic block (*Clinical Orthopaedics and Related Research).*

Laslett, M., McNair, P., Cadogan, A., Hing, W. (under review). Comparison of a novel direct measure of rapid pain intensity change to traditional serial 100mm VAS measurement of pain intensity (*Clinical Journal of Pain).*

Conference Presentations

Invited Presentations (plenary address)

Cadogan, A., Laslett, M., Hing, W., & McNair, P., Taylor, S. (2011). Subacromial impingement – what are we treating? NZ Manipulative Physiotherapists Association Biennial Conference, Rotorua, NZ.

Laslett, M., McNair, P., Cadogan, A., Hing, W. (2011). How to measure change in pain intensity: A novel and simple method. Paper presented at the Biennial Scientific

Conference of the New Zealand Manipulative Physiotherapists Association, Rotorua, August 2011.

Cadogan, A., Laslett, M., Hing, W., & McNair, P. (2010). Diagnosis of shoulder pain. Presented at the Nelson Shoulder Symposium, Nelson, New Zealand.

Cadogan, A., Hing, W., Laslett, M., & McNair, P, Mohammed, K. (2009). Clinical diagnosis of glenoid labrum pathology. Sports Medicine New Zealand Conference, Rotorua, NZ.

Published Abstracts

Cadogan, A., Laslett, M., Hing, W., & McNair, P. (2011). Can we diagnose subacromial pain using clinical examination and imaging findings? NZ Manipulative Physiotherapists Association Biennial Conference, Rotorua, NZ.

Laslett, M., McNair, P., Cadogan, A., Hing, W. (2011). How to measure change in pain intensity: A novel and simple method. Paper presented at the Biennial Scientific Conference of the New Zealand Manipulative Physiotherapists Association, Rotorua, August 2011.

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Williams, M. (2009c). Interexaminer reliability of orthopaedic special tests of the shoulder. Australian Physiotherapy Association Conference Week, Sydney, Australia.

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Williams, M. (2009d). Interexaminer reliability of orthopaedic tests used in assessment of the shoulder. NZ Manipulative Physiotherapists Association Biennial Conference, Rotorua, NZ.

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Williams, M. (2009e). Interexaminer reliability of range of motion, resisted muscle tests and associated clinical responses to physical examination tests of the shoulder. NZ Manipulative Physiotherapists Association Biennial Conference, Rotorua, NZ.

Papers in Published Proceedings

Cadogan, A., Laslett, M., Hing, W., & McNair, P. (2009a). Diagnostic accuracy of clinical examination findings for detecting pain arising from the subacromial bursa. Paper presented at the meeting of the Australian Physiotherapy Association Conference Week Sydney, Australia.

Cadogan, A., Laslett, M., Hing, W., & McNair, P. (2009b). Diagnostic accuracy of clinical tests for rotator cuff pathology. Paper presented at the Australian Physiotherapy Association Conference Week, Sydney, Australia.

Oral Presentations

Cadogan, A., Laslett, M., Hing, W., & McNair, P. (2010). Diagnosis of shoulder pain at primary care level: the research experience. Invited oral presentation: University of Otago, Christchurch School of Medicine, Christchurch, NZ.

Cadogan, A., Laslett, M., Hing, W., & McNair, P. (2010). Diagnosis of shoulder pain at primary care level: preliminary results. Invited oral presentation: Physiosouth, Christchurch, NZ.

Cadogan, A., Laslett, M., Hing, W., & McNair, P. (2010). Prevalence of shoulder pathology and correlation between imaging findings and response to diagnostic blocks. Christchurch Radiology Group, Christchurch, NZ.

Cadogan, A., Laslett, M., Hing, W., McNair, P, &Williams, M. (2008). Reliability of clinical tests of the shoulder. Active Health QEII, Christchurch, NZ.

Cadogan, A., Laslett, M., Hing, W., & McNair, P. (2008). Diagnostic accuracy of a clinical examination in determining the source of shoulder pain: Pilot study results. Active Health QEII, Christchurch, NZ.

Cadogan, A., Laslett, M., Hing, W., & McNair, P. (2008). Diagnostic accuracy & reliability of clinical tests of the shoulder: Research review. Barrington Physiotherapy Clinic, Christchurch, NZ.

ACKNOWLEDGEMENTS

ETHICS

Ethical approval for this study was granted on 9/5/2008 by the Ministry of Health Regional Ethics Committee (Upper South A). (Reference 08/02/013). Copies of ethical approval, participant information sheets and consent forms are included in Appendices 1-3 (p257-268).

ABSTRACT

Shoulder pain is a common and disabling complaint that is associated with high morbidity and significant associated health care costs. Shoulder pain is a common reason for primary care medical consultation, however the clinical diagnosis of shoulder pain is complicated by the similar presentations of different shoulder conditions and a lack of validated clinical examination tests and diagnostic criteria in primary care populations. Radiological imaging is also widely available and is increasingly being utilized to aid in the diagnostic process however the relevance of imaging to symptoms of shoulder pain remains unclear. The difficulties associated with the diagnosis of shoulder pain frequently result in delayed diagnosis and delays in the implementation of appropriate management. An improvement in the ability to accurately diagnose painful shoulder conditions would assist in optimising patient outcomes in primary health care.

The aim of this thesis was to evaluate the diagnostic accuracy of a clinical examination for identifying a predominant subacromial, acromioclavicular joint (ACJ) and glenohumeral joint (GHJ) pain source, and to assess the added value of diagnostic imaging findings for identifying symptomatic pathology affecting these structures. The diagnostic accuracy of clinical examination findings for detecting the presence of rotator cuff tears that may require early referral for specialist evaluation was also assessed. A review of the literature highlighted the poor specificity of many commonly used clinical tests, a lack of information regarding the relationship between imaged pathology and symptoms with little information to guide decisions regarding the use of diagnostic imaging investigations for shoulder pain.

This project consisted of a reliability study in which the reliability of clinical examination tests was first evaluated (Chapter 3), followed by a diagnostic accuracy study in which consecutive patients with shoulder pain were recruited from primary health care physiotherapy and medical practices. All participants received a clinical examination and a series of diagnostic imaging investigations (x-ray and diagnostic ultrasound scan) (index tests) followed by a diagnostic injection of local anaesthetic (diagnostic block) into the subacromial bursa (SAB) and ACJ (reference standard tests) (Chapter 4). Those not reporting at least 80% reduction in pain (positive anaesthetic response (PAR)) following the SAB or ACJ diagnostic block also received a GHJ diagnostic block performed as part of a magnetic resonance arthrogram (MRA) investigation. Results of the clinical examination and diagnostic imaging investigations

(index tests) were compared with results of the reference standard tests to estimate the ability of these clinical examination and imaging findings to accurately identify a predominant subacromial, ACJ or GHJ pain source and to detect the presence of rotator cuff tears.

Combinations of clinical features were identified with the ability to accurately rule-in a PAR following SAB and ACJ diagnostic block. When only a small number of these clinical features were present, confirmation of supraspinatus or ACJ pathology on ultrasound improved the ability to rule-in a PAR following SAB and ACJ diagnostic block respectively (Chapters 5 and 6). Overall the added diagnostic value of imaging findings for predicting an 80% PAR was limited due to the low prevalence of imaging findings, resulting in identification of only a small additional number of cases in whom a PAR could be predicted. Additional diagnostic investigations such as clinically-administered diagnostic injections of local anaesthetic may provide more information regarding the likelihood of a predominant subacromial or ACJ pain source in a larger proportion of patients. Analysis of diagnostic accuracy of clinical examination and imaging findings for predicting a PAR following GHJ diagnostic block was beyond the scope of this thesis but will be the subject of ongoing analysis. Clinical examination predictors of a large or multi-tendon rotator cuff were also identified that were able to accurately identify the presence of a large or multi-tendon rotator cuff tear that may require specialist evaluation (Chapter 7).

In conclusion, the ability to accurately diagnose painful subacromial and ACJ disorders in primary care begins with information gathered from the clinical examination however, for many patients, accurate diagnosis of these disorders may also require additional diagnostic investigations including diagnostic imaging or diagnostic injections. Combinations of clinical examination findings alone are likely to be sufficient to identify a large or multi-tendon rotator cuff tear that may require specialist evaluation. Results of this research may provide a framework that can be used by primary care practitioners to guide diagnostic processes for painful shoulder disorders, enabling more accurate and efficient identification of these conditions. This has the potential to reduce health care costs, reduce the burden on secondary care services, enable more timely application of appropriate treatment interventions and improve outcomes for patients suffering from shoulder pain.

THESIS ORGANISATION

The project undertaken for this thesis firstly involved a reliability study in which reliability of clinical examination procedures was evaluated, followed by a diagnostic accuracy study in which results of the clinical examination and diagnostic imaging investigations were compared with reference standard tests aimed at identifying sources of shoulder pain and pathology.

In Chapter 1 of the thesis, the background to the project is discussed and the aims of the thesis are presented. In Chapter 2, findings from a general review of the literature are reported that relate to the diagnostic accuracy of clinical examination findings for identifying sources of shoulder pain and pathology.

The methods and results of both the reliability study (Chapter 3) and the diagnostic accuracy study (Chapters 4 to 7) are then presented in a series of manuscripts. These manuscripts are arranged into the following chapters:

Chapter 3: Reliability of the Clinical Examination

Methods and results of the reliability study and details of subsequent methodological development of clinical examination testing procedures are presented in this chapter as a series of four manuscripts:

3.1 Reliability of a new hand-held dynamometer in measuring shoulder range of motion and strength (A. Cadogan, M. Laslett, W. Hing, P. McNair, & M. Williams, 2011a). Published in *Manual Therapy, 16*(1), 97-101.

3.2 Interexaminer reliability of orthopaedic special tests used in the assessment of shoulder pain (A. Cadogan, M. Laslett, W. A. Hing, P. J. McNair, & M. Williams, 2011b). Published in *Manual Therapy, 16*, 131-135.

3.3 Reliability of symptom responses associated with range of motion and resisted tests (unpublished results).

3.4 Methodological development for measures of range of motion and peak muscle force (unpublished results).

Chapter 4: Overview of Diagnostic Study Methods

The diagnostic accuracy study methods are explained in Chapter 4 in a published manuscript. In this manuscript some descriptive results are also presented relating to the diagnostic imaging (index tests) and diagnostic block procedures (reference standard tests):

4.1 A prospective study of shoulder pain in primary care: Prevalence of imaged pathology and response to guided diagnostic blocks (Cadogan, Laslett, Hing, McNair, & Coates, 2011). Published in *BMC Musculoskeletal Disorders, 12*, 119.

Chapter 5: Predictors of a Positive Response to Subacromial Bursa Diagnostic Block

The results of discrete analyses reporting the ability of clinical examination and imaging findings to predict a predominant subacromial pain source (positive response to subacromial bursa diagnostic block) are presented in Chapter 5:

5.1 Clinical predictors of a positive response to a guided diagnostic block into the subacromial bursa (Cadogan, Laslett, Hing, McNair, & Taylor). Manuscript submitted to *Journal of Rehabilitation Medicine*.

5.2 Added value of imaging findings for prediction a positive response to a guided subacromial bursa diagnostic block (unpublished results).

Chapter 6: Predictors of a Positive Response to Acromioclavicular Joint Diagnostic Block

Chapter 6 follows the same format as Chapter 5, and presents two manuscripts reporting results of the diagnostic accuracy of clinical examination and the added value of imaging findings for predicting a positive response following ACJ diagnostic block:

6.1 Clinical diagnosis of a positive response to a guided acromioclavicular joint diagnostic block (Cadogan, Laslett, Hing, McNair, & Taylor). Manuscript under review in *Clinical Orthopaedics and Related Research*.

6.2 Added value of imaging findings for prediction a positive response to a guided acromioclavicular joint diagnostic block (unpublished results).

Chapter 7: Clinical Diagnosis of Large Rotator Cuff Tears

In Chapter 7 results for the diagnostic accuracy of clinical examination findings for predicting a large or multi-tendon rotator cuff tears are presented:

7.1 Diagnostic accuracy of the clinical examination for predicting large rotator cuff tears (unpublished results).

Chapter 8: Summary, Discussion and Conclusions

Chapter 8 presents a summary of diagnostic accuracy study results, a discussion relating to overall findings from the study, clinical applications and conclusions. Limitations of the study and areas for future research are also identified in this chapter.

CHAPTER ONE

INTRODUCTION

1.1 The Problem

Shoulder pain is one of the most commonly reported musculoskeletal disorders, and is associated with significant disability, reduced health-related quality of life, depression and loss of functional independence (Croft, Pope, & Silman, 1996; Ostor, Richards, Prevost, Speed, & Hazleman, 2005). Musculoskeletal conditions, including shoulder pain, are placing an increasing burden on global health resources. With an increasing prevalence of musculoskeletal conditions in both developed and developing nations. The financial and healthcare burdens from these disorders are set to escalate dramatically (WHO Scientific Group, 2000).

Significance of Shoulder Pain

Painful shoulder conditions affect all age groups, levels of physical activity and types of occupational work. Shoulder pain can result from trauma (Mazzocca, Arciero, & Bicos, 2007), recreational or sports-related physical activity (Auge II & Fischer, 1998; Auvinen et al., 2007; Hawkins & Kennedy, 1980), occupational postures or tasks (Hagberg & Wegman, 1987) as well as sedentary activity (Auvinen et al., 2007). Shoulder pain is also associated with increasing age, cervicothoracic spine dysfunction and with certain medical conditions including metabolic disease (diabetes mellitus and thyroid abnormalities), rheumatologic conditions, osteoarthritis and stroke (Bjelle, 1989; Cole et al., 2009; Lindgren, Jonsson, Norrving, & Lindgren, 2007; Meislin, Sperling, & Stitik, 2005; Petersson, 1986; V. Wright & Haq, 1976).

Shoulder pain is reported to be the third most common site of musculoskeletal pain in the general population after lower back pain and knee pain (Urwin et al., 1998). Up to two thirds of the population may be affected by shoulder pain at some time, with reported lifetime prevalence rates for shoulder pain ranging between 10% and 67% (Luime, Koes, et al., 2004), similar to the reported lifetime prevalence rates for lower back pain (11% to 84%) (Walker, 2000). The prevalence of shoulder pain also increases with age, with a point prevalence of up to 34% in those over 65 years of age (Chakravarty, 2002).

Shoulder pain is a common reason for consultation in primary health care (Linsell et al., 2006; Ostor et al., 2005). Subacromial disorders (subacromial bursitis, rotator cuff disease, rotator cuff tears and 'impingement') are reported to be the most common disorders accounting for up to 85% of shoulder disorders seen in primary care, and acromioclavicular joint (ACJ) and glenohumeral joint (GHJ) disorders are also common (16% to 30% of shoulder pain) (Chard, Hazelman, Hazelman, King, & Reiss, 1990; Ostor et al., 2005; D. A. W. M van der Windt, Koes, De Jong, & Bouter, 1995; Vecchio, Kavanagh, Hazleman, & King, 1995). Incidence rates for shoulder pain consultations in primary care are reported to vary between 11.0 and 23.0/1000 person/years, equating to the average medical practice serving 2500 registrants conducting an average of 148 consultations per year for new episodes of shoulder pain (Bot et al., 2005; D. A. W. M van der Windt et al., 1995). Multiple consultations for shoulder pain are also common, with approximately half of patients consulting more than once for the same episode of shoulder pain (Linsell et al., 2006). Older patients are more likely to consult for longer with 18% of those aged over 60 years still consulting after 24-month follow-up, compared with 13% of those aged 40-59 years, and only 8.7% of those aged 18-39 years (Linsell et al., 2006).

Recovery from shoulder pain can be slow, and recurrence is common. Up to one quarter of patients presenting with a new episode of shoulder pain reported a previous episode affecting the same shoulder and only 21% of patients reported full resolution of shoulder pain symptoms after six months (Croft et al., 1996). A further 40% to 50% of patients were still experiencing ongoing pain and functional limitation after 1-2 years (Croft et al., 1996; D. A. W. M van der Windt et al., 1996; J. C Winters, Sobel, Groenier, Arendzen, & Meyboom - de Jong, 1999). Shoulder pain appears to be even more persistent in the elderly (>70 years) with most shoulder disorders continuing to cause symptoms after 3 years and several reports of symptom duration of up to 10 years (Chard et al., 1990).

The high prevalence of shoulder pain, high consultation rates and lengthy recovery times have resulted in a considerable increase in the cost of diagnosis and management of injury-related shoulder pain in New Zealand since 2000. In the 2010 financial year, the New Zealand Accident Compensation Corporation (ACC) received 108,700 claims for a new shoulder 'injury', with 32,000 ongoing claims amounting to NZD $106,000,000 for the 12-month period (Information Services ACC, 2011). Almost half this amount (NZD $43,000,000) was spent on medical treatment fees including

primary care practitioner consultations and diagnostic imaging costs. These figures do not include costs of publically funded and non-injury related shoulder pain and the actual cost is therefore likely to be much higher.

The accurate diagnosis of shoulder conditions is therefore important in managing cost and resource utilisation, and enabling application of appropriate treatment interventions that optimise recovery and treatment outcomes for patients with shoulder pain presenting to primary health care practitioners.

Diagnosis of Shoulder Pain in Primary Health Care

Importance of Diagnosis

In medical practice, diagnosis is considered the keystone of patient management. Diagnosis serves as a link between examination findings and interventions, providing the basis for rational decisions regarding treatment selection and patient management (Fritz & Wainner, 2001). The assumption is that accurate diagnosis of the condition will lead to more targeted treatment selection resulting in improved outcomes.

In musculoskeletal medicine it has been suggested that the exact localisation of the anatomical site of the lesion is a prerequisite for effective treatment (S. Green, Buchbinder, & Hetrick, 2007), with different management strategies warranted for different conditions (Baring, Emery, & Reilly, 2007). Conservative management including physiotherapy may be appropriate for some conditions (S. Green, Buchbinder, & Hetrick, 2003; J. C. Winters, Sobel, Groenier, Arendzen, & Meyboom-De Jong, 1997), injection of corticosteroid and/or drug therapy may be appropriate in the management of painful SAB, ACJ or GHJ conditions (Buchbinder, Green, & Youd, 2003; S Carette et al., 2003; Plafki, Steffen, E., & Wittenberg, 2000; D. A. W. M van der Windt et al., 1998), interventional radiology procedures such as fenestration or barbotage may be appropriate for calcific tendon lesions (Comfort & Arafiles, 1978) and surgical treatment options exist for other conditions such as glenoid labrum lesions and rotator cuff tears (Levy, Gardner, & Lemak, 1991; Stephen J. Snyder et al., 1991; Wasserlauf & Matava, 2003). Whether such interventions are appropriately applied depends upon the accurate diagnosis of the condition in the first instance.

Diagnostic Processes in Primary Health Care

In primary health care practice, the diagnostic process typically begins with a clinical examination that includes gathering information from both the patient history

and physical examination tests. This information is used to aid in formulating a diagnosis that is used to screen for (rule-out) specific conditions that may not be appropriate for management in primary care, or that may require concurrent medical management such as rheumatologic disease (Vecchio et al., 1995). Clinical examination tests are also used to aid in identifying (ruling-in) those patients with conditions for which specific management or treatment interventions are likely to be beneficial, or for whom additional medical, radiological or surgical investigations may be appropriate. Where the diagnosis remains unclear following the clinical examination, additional diagnostic tests such as radiological imaging investigations are commonly used to rule-in or rule-out specific shoulder disorders. However several issues surround the clinical and radiological diagnosis of shoulder pain and, for many patients for whom a definitive diagnosis cannot be reached, the diagnostic process culminates with a referral for orthopaedic consultation to assist in confirming the diagnosis.

Clinical examination.

The clinical examination in primary care typically involves the collection of patient history information and a physical examination aimed at determining the source of pain and presence of pathology. In the absence of a standardised approach to the clinical evaluation of the shoulder, clinicians currently rely upon a variety of non-standardised clinical tests, selected according to individual practitioner preference, training background, beliefs and experience. Many of the physical examination tests used by clinicians have demonstrated variable reliability and lack diagnostic validity in the primary care setting (Bohannon, 1999; Gajdosik & Bohannon, 1987; Hegedus et al., 2008; Johansson & Ivarson, 2009; Norregaard, Krogsgaard, Lorenzen, & Jensen, 2002).

The clinical diagnosis is also hampered by differences in patient interpretation of the location of "shoulder pain", and similar clinical presentations of different shoulder disorders (Ostor et al., 2005; Pope, Croft, Pritchard, & Silman, 1997). The complex regional anatomy involving intimate anatomical and functional relationships between various structures means multiple pathologies frequently coexist, and the ability to isolate specific structures with physical testing is limited. Consequently, many clinical tests lack specificity for detecting the specific shoulder pathology for which they are reportedly used (MacDonald, Clark, & Sutherland, 2000; Parentis, Glousman, Mohr, & Yocum, 2006; Park, Yokota, Gill, el Rassi, & McFarland, 2005; Walton et al., 2004). In addition, many previous diagnostic studies used surgery or diagnostic imaging as the reference standard procedure for identifying the presence of pathology. Reference

standard procedures that are based upon visualisation of pathology during surgery or on radiological imaging do not take into account whether the observed pathology was the source of symptoms.

Diagnostic imaging.

The inability to confidently locate the tissue origin of pain following the clinical examination frequently results in referral for diagnostic imaging investigations to aid diagnosis. Radiographs (x-ray) and diagnostic ultrasound imaging are the most widely available diagnostic imaging investigations available to primary care practitioners in New Zealand, and magnetic resonance imaging (MRI) is also available through high tech imaging referral pathways (Arrol et al., 2004). Difficulties associated with clinical diagnosis of shoulder pain are thought to be behind the increasing use of diagnostic imaging by primary care practitioners (Awerbuch, 2008). An Australian study reported that approximately 70% of patients with shoulder pain are referred for diagnostic imaging investigations at their first consultation, with a further 32% referred for imaging investigations at a subsequent visit (Broadhurst, Gialamas, McElroy, & Beilby, 2004). However, a limited number of clinical guidelines for the use of diagnostic imaging in primary care are available, and all appear to be based upon limited literary evidence and provide many non-specific guidelines to accommodate different clinical presentations and health care funding regulations (Arrol et al., 2004; Bussières, Peterson, & Taylor, 2008).

While imaging may aid identification of tissue pathology, interpretation of pathological imaging findings with respect to their contribution to patient symptoms is often complicated by the presence of anatomic variants and the high prevalence of asymptomatic pathology especially in ageing populations (De Maeseneer et al., 2000; Milgrom, Schaffler, Gilbert, & van Holsbeeck, 1995; Sammarco, 2000; Shubin Stein, Wiater, Pfaff, Bigliani, & Levine, 2001). With little available information regarding the prevalence of imaged pathology and the relationship between imaging findings and symptoms in those suffering from a current episode of shoulder pain, interpretation of the relevance of any reported imaging findings is difficult. Misinterpretation of imaging information may lead to incorrect diagnosis of symptom origin and inappropriate treatment pathways that, in some cases, may culminate in surgical intervention. The ad hoc use of diagnostic imaging, difficulties with interpretation of their results and lack of clear guidelines may add to the diagnostic confusion in many cases resulting in unnecessary costs to the health system and inappropriate management.

Referral to secondary care services.

Referral of primary care patients for consultation with rheumatology or orthopaedic specialists is also common. Reports suggest up to 41% of patients presenting to a primary care physician with shoulder pain may be referred for specialist assessment within one year of initial presentation (Solomon et al., 2001; D. A. W. M van der Windt et al., 1995). One study reported that referrals for orthopaedic evaluation demonstrated a universal lack of clinical examination or other diagnostic information, yet 99% of patients referred had received radiological investigation (Johal et al., 2008). An "uncertain diagnosis" appears to be a common reason for referral to secondary care services and evidence suggests up to 30% of referrals to specialists, including musculoskeletal specialists may be potentially avoidable or inappropriate (Donohoe et al., 1999).

The increasing use and apparent reliance upon radiological imaging by general practitioners, and the high rate of referral to orthopaedic care in some countries may reflect patient expectation, a lack of confidence in the clinical diagnosis of shoulder pain, or both. Regardless of the reason, many of these referrals may be unnecessary, resulting in increased health care spending, and delayed diagnosis and treatment while patients endure lengthy delays on imaging or specialist waiting lists. The ability to accurately and efficiently identify those patients for whom radiological imaging or specialist assessment is needed would assist in more accurate identification of the source of symptoms, and may reduce demand for services, reduce waiting times, and associated costs to the health care system.

Diagnostic Classification of Shoulder Pain

It appears there are no universally accepted criteria for the diagnostic classification of shoulder pain in primary health care. Guidelines published by the Accident Compensation Corporation of New Zealand identified "a dearth of sound evidence to guide the diagnosis and management of shoulder injuries" (pg 1) (Arrol et al., 2004).

Traditional classification of shoulder pain in medical practice is based upon the concepts of James Cyriax's selective tissue tension model (Cyriax, 1978). This system is based upon a combination of history and physical examination findings involving tests reported to selectively tension contractile or inert structures, with concurrent assessment of joint 'end-feels'. However, evidence for the criterion validity of this

method of assessment with regard to identification of specific pathology appears to be lacking. Other proposed classification criteria are population-specific including predominantly occupational settings (Harrington, Carter, Birrell, & Gompertz, 1998) or hospital rheumatology clinics (K. Palmer et al., 2000), with only one developed by multidisciplinary Delphi consensus (Davis, 1998). However reliability and validity was not assessed in that study.

Inconsistent diagnostic terminology also hampers consistent diagnostic labelling of shoulder pain. Some diagnostic terms relate to symptoms and signs derived from the clinical examination such as 'painful arc syndrome', a clinical hypothesis such as 'impingement syndrome' or 'frozen shoulder', or to a specific pathoanatomic diagnosis such as subacromial bursitis, tendinosis or rotator cuff tear. The terminology used also varies according to the radiological information available at the time, and also varies among medical professionals including general medical practitioners, physiotherapists, rheumatologists, musculoskeletal medicine, occupational health and sports medicine physicians, and orthopaedic surgeons. Such inconsistencies affect the clinical diagnosis, resulting in subsequent ad-hoc and heterogeneous treatment pathways.

The apparent lack of validated diagnostic criteria is also negatively affecting research advances into treatment interventions for shoulder pain. Outcomes of intervention trials are affected by poorly defined, non-validated or overlapping case definitions of patient groups in the first instance, frequently resulting in many inconclusive findings. Results of these trials are often misinterpreted as evidence of 'no effect', and as reflecting a general homogeneity of shoulder pain, thus contributing to a proliferation of the term "non-specific shoulder pain" (Helliwell, Bennett, Littlejohn, Muirden, & Wigley, 2003; Miranda, Viikari-Juntura, Heistaro, Heliovaara, & Riihimaki, 2005). This has prompted some to question whether a diagnosis is even required and whether physiotherapists should be treating shoulder pain at all (Smidt & Green, 2003). The lack of demonstrable treatment effects in previous intervention trials poses the risk of influencing policy change with regard to funding of specific interventions for specific disorders. Validated diagnostic criteria are required as a start-point from which potentially beneficial interventions can be tested.

Summary of the Problem

Shoulder pain is a common and disabling complaint seen in primary health care that represents significant cost in both diagnosis and management to consumers and funders of health care services. Current diagnostic processes for shoulder pain in primary care are inconsistent and based upon a lack of evidence for the diagnostic validity of many commonly used clinical examination and diagnostic imaging tests in the primary health care setting. Much of the evidence available from other settings is based upon comparison with reference standards which assume that observed pathology is the source of symptoms despite evidence reporting high prevalence of such pathology in asymptomatic individuals. The lack of ability to accurately diagnose shoulder pain in primary health care results in many patients experiencing delays in receiving a diagnosis, with subsequent delays in implementation of appropriate management.

The increasing trend for management of musculoskeletal conditions in primary health care means practitioners in this setting need to be able to accurately diagnose shoulder pain to enable implementation of timely and appropriate management for specific shoulder conditions. Evaluation of the diagnostic accuracy of clinical examination and diagnostic imaging tests, and identification of diagnostic tests that are of most value for identifying specific sources of shoulder pain and pathology is required in primary health care.

1.2 Thesis Aims and Objectives

The aims of the thesis are as follows:

General Aim

To estimate the diagnostic accuracy of a standardised clinical examination for identifying common sources of shoulder pain and specific shoulder pathology in primary care using available reference standards, and to assess the contribution of diagnostic imaging findings to the diagnosis of these conditions.

Specific Aims

1. To evaluate the reliability of clinical examination tests included in a standardised clinical examination.

2. To estimate the diagnostic accuracy of clinical examination findings for identifying subacromial, ACJ and GHJ sources of shoulder pain defined by a positive response to diagnostic block.

3. To evaluate the prevalence of imaged pathology and assess the relationship between imaging findings and anaesthetic responses to subacromial bursa (SAB), ACJ and GHJ diagnostic blocks.

4. To evaluate the added diagnostic value of imaging findings for predicting a positive response to SAB, ACJ and GHJ diagnostic blocks.

5. To estimate the diagnostic accuracy of clinical examination findings for identifying large rotator cuff tears.

Significance of the Research

Results will be of particular relevance to primary health care practitioners (physiotherapists and general medical practitioners) by providing information regarding the clinical examination features that are of most value for accurately identifying specific sources of shoulder pain, and pathology that may require early referral for specialist assessment. Results may also provide information regarding the symptomatic relevance of imaged shoulder pathology, and the relative value of diagnostic imaging findings for aiding identification of specific sources of shoulder pain.

The potential benefits of improvements in diagnostic accuracy of these conditions include more timely application of appropriate treatment interventions or referral for specialist evaluation with the potential for improved treatment outcomes. Refined diagnostic processes may also reduce the costs associated with diagnosis and management of shoulder pain including primary care consultations and diagnostic imaging. An improved ability to diagnose shoulder pain in primary care settings may also reduce the burden on secondary care services for patients with 'undiagnosed' shoulder pain. Identification of the best predictors of specific shoulder conditions may also aid refinement of diagnostic classification criteria for shoulder pain from which the effectiveness of various treatment interventions can be tested.

CHAPTER TWO

LITERATURE REVIEW

The clinical diagnosis of shoulder pain is challenging, with a lack of clear diagnostic criteria for painful shoulder conditions, and no clear guidelines for the use of diagnostic imaging. An improvement in the ability to accurately identify primary sources of shoulder pain and pathology is urgently required in primary health care. The aim of the literature review was to summarise what is currently known regarding the ability of clinical tests to diagnose specific shoulder disorders, and evaluate available evidence for the validity of reference standard tests used to detect these disorders.

2.1. Literature Review Methods

In accordance with the aims of this thesis, the following questions were identified:

Literature Review Questions

1. **What are the common causes of shoulder pain in primary care?**

   Evidence was reviewed regarding epidemiological aspects of shoulder pain with specific attention to those conditions commonly presenting to primary care practitioners.

2. **What evidence exists for the diagnostic accuracy of clinical examination components in identifying these conditions?**

   Evidence was reviewed regarding the diagnostic accuracy of both history and physical examination aspects of the clinical examination with specific attention to the reported ability of these tests to detect specific sources of shoulder pain or specific pathology commonly seen in primary care.

3. **What evidence exists for the relevance of diagnostic imaging findings in identifying symptomatic shoulder pathology?**

   Evidence was reviewed relating to the association between imaging findings from investigations that are available in primary health care, and the symptoms of shoulder conditions that are commonly seen in this health care setting.

4.   **What evidence exists for the validity of available reference standards for identifying sources of shoulder pain and pathology?**

Information was summarised regarding the use, and the validity of available reference standard procedures for detecting sources of musculoskeletal pain with particular attention given to their ability to detect specific sources of shoulder pain. Information regarding the validity of available imaging procedures to detect specific shoulder pathology was also evaluated.

Literature Search Strategy

A general literature search was undertaken using the following electronic databases: MEDLINE via PubMed, MEDLINE via OVID, MEDLINE via EBSCO, CINAHL, SPORT-Discus, AMED, PEDro, ProQuest 5000, Health and Psychosocial Instruments and SCOPUS. The search was limited to articles involving human participants published in the English language prior to July 2009. Article titles and abstracts were screened for relevance and the bibliographies of key articles were reviewed to identify other relevant articles which were entered into the SCOPUS database. A full systematic review of each area was beyond the scope of this thesis. The key concepts and specific search terms relating to each of the literature review questions are presented in the following sections.

2.2. Shoulder Pain in Primary Care

Search Terms

The concepts and search terms in Table 2.1 were used to identify the most common causes of shoulder pain in primary care.

Literature Search Results

A limited number of studies were identified reporting prevalence estimates for specific shoulder disorders in primary care and community settings (Chard et al., 1990; Ostor et al., 2005; Schaardenburg, van den Brande, Ligthart, Breedveld, & Hazes, 1994; D. A. W. M van der Windt et al., 1995). Diagnosis of these disorders in all studies was

Table 2.1. *Search Terms for Shoulder Pain in Primary Care*

Concept	Combine search terms	Subject headings and search terms	Results
Shoulder pain	OR	Shoulder.sh exp Shoulder/ OR exp Shoulder Joint/ OR exp Shoulder Pain/ OR exp Shoulder Impingement Syndrome/OR exp Acromioclavicular joint/ OR or exp Rotator Cuff/ supraspinatus.mp OR infraspinatus.mp OR subscapularis.mp OR "teres minor".mp labr* AND shoulder.mp OR (glenoid ADJ3 lab*) bicep*.mp OR "bicep* brachii".mp OR "bicep* tend*".mp OR (bicep* ADJ5 sheath).mp (burs* OR tend*) AND shoulder.mp ("glenohumeral joint" OR shoulder) ADJ3 arthr*.mp (adhesive ADJ2 capsul*).mp exp Joint capsule/ OR 'frozen shoulder'.mp (shoulder adj2 instability).mp	47645
Prevalence	OR	exp Cross sectional studies/ OR cross sectional studies.mp epidemiology.mp OR prevalen*.mp incidence.tw retrospective*.mp OR prospective*.mp OR survey.mp OR questionnaire.mp	986340
Primary care	OR	exp Primary health care/ OR primary health care.mp (primary ADJ3 care).mp OR community.mp OR "general practi*".mp	387972
Concepts	AND	Shoulder pain Prevalence Primary Care	173
		Limit to english language, humans, years 1940-2009, remove duplicates.	142
		Title, abstract and bibliography search	24

based upon clinical test criteria, with the majority of studies reporting prevalence of diagnostic categories that included rotator cuff tendinopathy, rotator cuff tears, ACJ pathology, adhesive capsulitis (Chard et al., 1990; Ostor et al., 2005; D. A. W. M van der Windt et al., 1995; Vecchio et al., 1995) and primary glenohumeral joint osteoarthritis (Chard et al., 1990). The clinical diagnosis was not confirmed by radiological or surgical investigation in any of these studies. In addition, studies varied according to location (country), diagnostic criteria and terminology, study population (community, specific age groups, general medical practice), study design (cross-sectional, survey, retrospective and prospective) and the method of reporting incidence and prevalence making direct comparisons between studies difficult. Studies reporting the prevalence of shoulder conditions in primary care or general practice were extensively reviewed and findings from these studies relating to the prevalence of specific shoulder disorders seen in primary care are summarised in this section. Results are grouped according to the most commonly reported diagnostic categories.

Prevalence of Subacromial Pathology

Subacromial disorders, including subacromial bursitis, rotator cuff disease, rotator cuff tear and "impingement" are reported to be the most common shoulder conditions seen in primary care, with rotator cuff tears and impingement reported to respectively account for up to 85% and 74% of shoulder pain seen in this setting (Chard et al., 1990; Ostor et al., 2005; D. A. W. M van der Windt et al., 1996; Vecchio et al., 1995). In two of these studies, subacromial pathology was defined using case definitions reported by Cyriax, however each measured only incidence or prevalence, and diagnostic labels varied in these studies despite reported use of similar testing procedures (Ostor et al., 2005; D. A. W. M van der Windt et al., 1995). Two other studies used a combination of diagnostic criteria and reported between 65% to 70% prevalence of subacromial pathology (predominantly rotator cuff pathology) however prevalence could only be directly compared in hospital populations in these studies (Chard et al., 1990; Vecchio et al., 1995). All studies used the Cyriax classification for rotator cuff tendinitis, however this is based upon results of resisted tests which have been shown to be unreliable for shoulder conditions (K. W. Hayes & Petersen, 2003). Radiological imaging was also used in several patients in one study to help clarify the diagnosis, however application of imaging procedures was not performed in a systematic manner (Vecchio et al., 1995).

Diagnostic criteria of 'impingement', defined by the presence of a painful arc during abduction, also overlapped with those of 'rotator cuff tendinitis' and rotator cuff tear (Chard et al., 1990; Ostor et al., 2005). Although not specifically stated, it appears pain with resisted testing (rotator cuff tendinitis) and inability to raise the arm or significant weakness (rotator cuff tear) differentiated 'impingement' from either of these two pathologies (Chard et al., 1990; Ostor et al., 2005). It was not reported however, how many participants had both and would therefore have satisfied more than one of the diagnostic criteria.

Overlapping diagnostic criteria, differences between diagnostic criteria and diagnostic labelling and sampling in these studies mean that while it can be concluded that symptoms that are assumed to be of subacromial origin are the most common shoulder condition seen in primary care, lack of reconciliation between clinical and imaging evidence, particularly with regard to the diagnosis of 'impingement' means the exact prevalence of subacromial pathology in a primary care population remains unknown.

Prevalence of Acromioclavicular Joint Disorders

Disorders of the acromioclavicular joint (ACJ) are reported to account for between 5% (D. A. W. M van der Windt et al., 1995) and 31% (Ostor et al., 2005) of all shoulder pain seen in primary care. Compared with other shoulder pathology, painful ACJ conditions were the second most common clinically-diagnosed shoulder condition in primary care (10% to 31% of all shoulder conditions) (Chard et al., 1990; Ostor et al., 2005; Vecchio et al., 1995). Diagnostic criteria and labelling were more consistent for ACJ pathology, with most studies using pain localized to the ACJ or C4 dermatome, localized tenderness and symptom provocation with stress testing as the diagnostic criteria. However no confirmation of painful ACJ pathology was obtained using either radiological imaging or diagnostic injection of local anaesthetic in these studies.

Prevalence of Glenohumeral Joint Disorders

Glenohumeral joint disorders constitute between 16% (Ostor et al., 2005) and 21% (D. A. W. M van der Windt et al., 1995) of shoulder disorders seen in primary care. In one study involving patients recruited from general practice, 21% were classified as having capsular syndrome (capsular pattern of motion restriction), representing the second most common diagnostic category in this study following subacromial conditions (D. A. W. M van der Windt et al., 1995).

The most commonly reported glenohumeral joint disorders are arthrosis and adhesive capsulitis ('frozen shoulder') (Chard et al., 1990; Ostor et al., 2005; Schaardenburg et al., 1994; D. A. W. M van der Windt et al., 1995). Adhesive capsulitis was reported to affect 16% of those with shoulder pain recruited from primary care practice (Ostor et al., 2005), and 17% of those over the age of 85 years (Schaardenburg et al., 1994). Glenohumeral joint osteoarthritis was more common in those over the age of 60 years, accounting for 2-3% of shoulder pain in community populations (Chard et al., 1990; Schaardenburg et al., 1994; Vecchio et al., 1995), and was present in approximately 5% of those with existing shoulder disease (Y. Nakagawa, Hyakuna, Otani, Hashitani, & Nakamura, 1999).

The diagnosis of GHJ disorders in these studies was based upon clinical evidence of loss of passive range of motion (ROM). As a result, the prevalence of other GHJ disorders including labral tears, isolated chondral lesions or rotator interval pathology that may have specific management implications in primary care practice was not reported in any studies. The prevalence of these conditions in a primary care population therefore remains unknown.

Prevalence of Shoulder Pathology on Radiological Imaging

In all the studies reviewed, the diagnosis of shoulder disorders was based upon results of clinical tests alone, and radiological investigations were not conducted to verify the diagnosis in any of the studies. While a number of studies were identified reporting the prevalence of radiological imaging findings in the general population (Milgrom et al., 1995) and asymptomatic samples (Moosmayer, Smith, Tariq, & Larmo, 2009), there is a lack of evidence regarding the prevalence of pathology diagnosed by radiological imaging in primary care patient populations. The prevalence of diagnostic categories according to imaged pathology in those suffering a current episode of shoulder pain therefore remains unknown.

Key Findings

- The reported prevalence of shoulder conditions in primary care is based predominantly upon clinical test criteria that vary between studies.
- Subacromial disorders are the most commonly reported causes of shoulder pain in primary care according to clinical test criteria however comparisons between studies is difficult due to variable diagnostic criteria.
- Acromioclavicular joint and glenohumeral joint disorders (adhesive capsulitis) are also common conditions seen in primary care and glenohumeral joint arthritis is also seen.
- The prevalence of shoulder pathology diagnosed by radiological imaging in a primary care cohort has not been reported.

2.3. Diagnostic Accuracy of the Clinical Examination

The literature was reviewed to evaluate available evidence for ability of clinical examination tests to accurately diagnose subacromial, ACJ and GHJ disorders.

Search Terms

The concepts and search terms listed in Table 2.2 were used in the literature search for studies reporting the diagnostic accuracy of clinical examination tests for subacromial, ACJ and GHJ disorders.

Table 2.2. *Search Terms for Diagnostic Accuracy of the Clinical Examination*

Concept	Combine search terms	Subject headings and search terms	Results
Shoulder pain		See Table 2.1.	47645
Clinical examination	OR	exp medical history taking/ OR medical history taking.mp exp physical examination/ OR physical examination.mp history.mp OR subjective exam* objective exam* clinical examination.tw	911210
Diagnostic accuracy	OR	exp Sensitivity and Specificity/ OR specificity.mp sensitiv*.mp. di.fs. valid*.tw OR accura*.tw OR specific*.tw OR "likelihood ratio".tw reliab*.mp OR varia*.mp OR agreement.mp	4151780
Concept	AND	Shoulder pain Clinical examination Diagnostic accuracy	2559
		Limit to English language, humans, years 1940-2009, remove duplicates.	218
		Title, abstract and bibliography search	109

Literature Search Results

Eighty five original diagnostic accuracy studies in which the diagnostic accuracy of clinical examination components was evaluated for identifying subacromial, ACJ or GHJ conditions. The majority of studies investigated the diagnostic accuracy of individual physical examination tests for identifying specific shoulder conditions, with several studies also reporting the diagnostic accuracy of combinations of clinical examination tests for detecting the presence of these disorders. Twenty four review articles, including eleven systematic reviews were also identified that critically appraised individual diagnostic accuracy studies reporting the diagnostic accuracy of physical examination tests for specific shoulder disorders. A number of methodological issues relating to the quality of diagnostic accuracy studies were also identified.

These literature review findings are grouped into sections below that summarise diagnostic accuracy results of individual clinical examination tests, combinations of clinical examination tests, and results of systematic reviews of diagnostic accuracy studies for subacromial, ACJ and GHJ pathology. Methodological issues relating to these studies are also summarised.

Diagnostic Accuracy of Individual Clinical Examination Tests

The majority of original diagnostic accuracy studies investigated the diagnostic accuracy of isolated physical examination (orthopaedic) tests (index tests) for identifying specific shoulder pathology by comparing results with surgical findings (reference standard procedure). A small number of studies estimated diagnostic accuracy of physical examination tests for identifying the source of pain, using a positive response to injection of local anaesthetic into either the subacromial space or ACJ as the reference standard (Calis et al., 2000; Chronopoulos, Kim, Park, Ashenbrenner, & McFarland, 2004; Walton et al., 2004). These studies were almost exclusively conducted in orthopaedic (secondary care) settings, and methodological flaws were common. With few exceptions, studies that investigated the diagnostic accuracy of physical examination tests tended to report either high sensitivity, or high specificity but not both. Few studies investigated the accuracy of patient history variables, or other aspects of the physical examination including ROM tests and resisted tests for identifying specific shoulder conditions.

Subacromial pathology

Few studies were identified in which the diagnostic accuracy of patient history variables for detecting subacromial pathology was reported with the majority of studies reporting the ability of physical examination tests to identify subacromial impingement and/or rotator cuff tears.

Subacromial impingement.

Subacromial impingement in the studies reviewed was generally defined according to the Neer classification of Stage I-III impingement, including subacromial bursitis, rotator cuff tendon pathology and rotator cuff tears (Neer, 1983). Clusters of clinical examination findings that included history variables (age >39 years and a history of shoulder "popping or clicking") combined with a painful arc during abduction were identified in one study as the best predictors of supraspinatus pathology diagnosed by ultrasound scan when all three were present (positive likelihood ratio (+LR) 32.2) (Chew, Pua, Chin, Clarke, & Wong, 2004) however these results have not been prospectively validated.

Of the physical examination tests used in individual diagnostic accuracy studies, consistently high sensitivity was reported for the Hawkins-Kennedy test for subacromial impingement (0.72 to 0.95)(Calis et al., 2000; Leroux, Thomas, Bonnel, & Blotman, 1995; MacDonald et al., 2000). Sensitivities for other subacromial impingement tests

were generally low to moderate and ranging from 0.08 (drop arm test) (Calis et al., 2000) to 0.82 (cross-body adduction test) (Calis et al., 2000). With the exception of the Hawkins-Kennedy test (Calis et al., 2000; MacDonald et al., 2000), no other test demonstrated sensitivity exceeding 0.90 for subacromial impingement pathology. The reported specificities were generally higher than sensitivities, ranging from 0.31 (cross-body adduction test) (Calis et al., 2000) to 1.00 (drop-arm test) (Calis et al., 2000), and specificity of the supraspinatus/empty can and infraspinatus tests were also reported to exceed 0.90 (Park et al., 2005).

The range of pathologies included in the diagnostic category of subacromial impingement (bursitis to full thickness rotator cuff tear) (Neer, 1983) are likely to exhibit some difference in clinical presentation (early stage bursitis compared with a full-thickness rotator cuff tear), and also have different management pathways. Therefore the diagnostic value of identifying 'subacromial impingement' (any stage) is questionable, and the range of pathologies included in the subacromial impingement outcome variable may explain the variable diagnostic accuracy reported in these studies. One study was identified specifically investigating the diagnostic value of physical examination tests for detecting different degrees of subacromial impingement pathology (Park et al., 2005). Results of this study found the diagnostic accuracy of the physical examination tests varied according to the stage of impingement pathology. The sensitivity of impingement tests was higher for subacromial bursitis (0.76 to 0.86) than for a full-thickness rotator cuff tear (0.60 to 0.76), and the specificity of specific muscle tests (drop-arm test and infraspinatus test) was higher for a full thickness tear (0.84 to 0.88) compared with the diagnosis of subacromial bursitis (0.69 to 0.77). However confidence intervals for the diagnostic accuracy estimates were not reported in this study (Park et al., 2005).

In contrast to all other studies in which surgery was used as the reference standard procedure for identifying subacromial impingement, Calis et al. (2000) used a subacromial injection of local anaesthetic and MRI (Calis et al., 2000). However the subacromial injection was performed without imaging guidance and the possibility that other 'non-impingement' structures were anaesthetised cannot be excluded. These studies included a number of biases including selection bias (non-consecutive patients), partial verification bias (not all participants received the reference standard) and lack of blinding of examiners.

Internal impingement.

One study was identified that specifically investigated that "posterior impingement sign" for identifying an articular surface rotator cuff, or posterior glenoid labrum tear (internal impingement) in a group of athletes with posterior shoulder pain (Meister, Buckley, & Batts, 2004). Sensitivity and specificity for the test were reported respectively as 0.76 and 0.85, but when only non-traumatic shoulder pain was analysed, sensitivity and specificity improved to 0.95 and 1.00 respectively. This represents a specific subgroup of the population and the prevalence of these pathologies in a primary care cohort is likely to differ limiting the applicability of these results to this population.

One other study investigated the ability of the internal rotation resistance test to differentiate between non-outlet (internal) impingement and outlet (subacromial) impingement (Zaslav, 2001). The authors reported both high sensitivity (0.88) and specificity (0.96) for this test, which was evaluated in patients with a positive Neer test who had failed conservative management and subsequently underwent arthroscopic investigation. Methodological weaknesses, including selection bias, non-standardised index test procedure, variable time-frames between index and reference standard procedures, lack of description of blinding processes and no reported confidence intervals hamper confidence in these results.

Rotator cuff integrity.

The relationship between patient history variables including age, pain location, night pain and history of trauma and a tear of the rotator cuff was reported in a small number of studies, with only one study identified in which the diagnostic accuracy statistics of these variables was reported (Litaker, Pioro, El Bilbeisi, & Brems, 2000). A history of trauma resulted in sensitivity and specificity respectively of 0.36 and 0.73 for a rotator cuff tear (Litaker et al., 2000). This study included patients with a suspected rotator cuff tear awaiting arthroscopy recruited from an orthopaedic waiting list and differences in prevalence of this condition in primary care, and lack of the confirmed presence of a cuff tear limit the generalisation of these results to primary health care settings. Older age (>65 years) (adjusted odds ratio (AOR 4.05)) and the presence of night pain (AOR 2.61) were identified as the best predictors of a rotator cuff tear in patients suspected of having this condition who were referred for arthrography (Litaker et al., 2000). Pain located in the anterior and lateral aspect of the shoulder was also associated with supraspinatus tears ($p<0.05$) in another study (Itoi, Minagawa, Yamamoto, Seki, & Abe, 2006).

From the physical examination, "lag" signs demonstrated consistently higher levels of diagnostic accuracy for rotator cuff tears than any other tests for any other shoulder pathology. The internal rotation lag sign for a partial or complete subscapularis tear (100% sensitivity) (Miller, Forrester, & Lewis, 2008a), lift-off (100% specificity) and belly press tests (98% specificity) for a subscapularis tear, external lag sign (94% sensitivity; 94% to 100% specificity) and drop arm test for supraspinatus or infraspinatus tear) (100% specificity) demonstrated consistently high levels of diagnostic accuracy (Barth, Burkhart, & De Beer, 2006; Castoldi, Blonna, & Hertel, 2009; Hertel, Ballmer, Lambert, & Berber, 1996; Miller, Forrester, & Lewis, 2008b). Only one study was identified investigating the relationship between strength and size of surgically identified rotator cuff tears (McCabe et al., 2005). Strength deficits of more than 50% at 10° of abduction were associated with a large rotator cuff tear ($p<0.001$) however, diagnostic accuracy was not reported.

Long head of biceps tendon.

A small number of studies investigated the diagnostic accuracy of physical examination tests for disease of the long head of biceps tendon (Ardic et al., 2006; Holtby & Razmjou, 2004; Leroux et al., 1995). The palm-up/Speed's test demonstrated limited diagnostic accuracy for biceps tendon pathology (sensitivity 0.32 to 0.69; specificity 0.35 to 0.75) in these studies. Yergason's test (sensitivity 0.43; specificity 0.79) also demonstrated marginal diagnostic specificity for biceps tendon pathology (Holtby & Razmjou, 2004).

Acromioclavicular joint pathology.

Only three studies were identified investigating the accuracy of physical examination tests for the diagnosis of ACJ pain defined by a positive response to diagnostic injection (Chronopoulos et al., 2004; Walton et al., 2004) or pathology identified on imaging (O'Brien, Pagnani, Fealy, McGlynn, & Wilson, 1998). Highest sensitivities were reported for the O'Brien's test (0.93) (O'Brien et al., 1998) and pain with palpation of the ACJ (0.96) (Walton et al., 2004). High specificity (0.90 to 0.96) was also reported for the O'Brien's test in all three studies. In the two studies of highest quality design, pain with palpation of the ACJ was the most sensitive test (0.96) and the O'Brien's test the most specific (0.90) for disorders of the ACJ (Walton et al., 2004). High diagnostic accuracy results reported by O'Brien et al., (1998), have not been independently verified to date.

Glenohumeral joint disorders

Adhesive capsulitis.

History variables relating to pain (strong component of night pain, pain with rapid or unguarded movement, discomfort lying on the affected shoulder, and pain easily aggravated by movement) were identified by expert consensus as important features of identifying early-stage adhesive capsulitis although prospective validation of these features was still pending at the time of commencing this study (Walmsley, Rivett, & Osmotherly, 2009).

Glenoid labrum tears.

A small number of studies investigated the association between patient history variables and a glenoid labrum tear. Higher levels of pain (r=0.8, *p*=0.000) and disability (r=0.6, *p*=0.020) were associated with glenoid labrum tears (Ardic et al., 2006), however diagnostic accuracy was not specifically reported. Combining a history of "popping, clicking or catching" in the shoulder with two positive physical examination tests (crank test and anterior slide test) improved the specificity for a glenoid labrum tear (0.82 to 1.00) compared with combinations of physical examination tests alone (0.64 to 1.00) (Walsworth, Doukas, Murphy, Mielcarek, & Michener, 2008). The inclusion of "popping, clicking or catching" among one of the required criteria increased sensitivity for a glenoid labrum tear (0.82 to 0.95) compared with criteria that did not include the history variable (0.70 to 0.82) (Walsworth et al., 2008). A history of trauma was also investigated for its diagnostic value for a glenoid labrum tear but sensitivity (0.50) and specificity (0.36) were low (Walsworth et al., 2008). Participants in this study had all failed conservative management programmes, and the prevalence of labral tears was high (85%). The spectrum of disease in this study is unlikely to be applicable to primary care populations.

Many studies were identified in which the diagnostic accuracy of physical examination tests for various labral pathology including superior labrum anterior-posterior (SLAP) lesions was investigated (Bennett, 1998; Guanche & Jones, 2003; Holtby & Razmjou, 2004; S. Kim, Ha, & Han, 1999; S. H. Kim, Ha, Ahn, Kim, & Choi, 2001; Y. S. Kim et al., 2007; McFarland, Kim, & Savino, 2002; Mimori, Muneta, Nakagawa, & Shinomiya, 1999; Myers, Zemanovic, & Andrews, 2005; S. Nakagawa et al., 2005; O'Brien et al., 1998; Parentis et al., 2006), posterior glenoid labral lesions (Heyworth & Williams, 2009; S. H. Kim, Park, Jeong, & Shin, 2005; S. H. Kim, Park, & Oh, 2004; Meister et al., 2004; Zaslav, 2001). The ability of clinical examination tests

to identify any type of labral pathology was also investigated by several authors (Guanche & Jones, 2003; Liu, Henry, Nuccion, S, & Dorey, 1996; Stetson & Templin, 2002).

The O'Brien's (active compression) test demonstrated the most consistent levels of sensitivity for a SLAP lesion (0.47 to 0.99) (McFarland et al., 2002; O'Brien et al., 1998), with the Jobe relocation test (specificity 0.31 to 0.98) and the anterior slide test (specificity 0.67 to 0.93) demonstrating the most consistently high levels of specificity of all tests identified (Guanche & Jones, 2003; S. Nakagawa et al., 2005; Parentis et al., 2006). High levels of both sensitivity and specificity were also identified for the biceps load I (sensitivity 0.91; specificity 0.97) and biceps load II tests (sensitivity 0.90; specificity 0.97) for a surgically identified SLAP lesion (S. Kim et al., 1999; S. H. Kim et al., 2001). However tests demonstrating the highest reported levels of diagnostic accuracy in these studies were conducted by the same person who developed the test and details of blinding of the examiner were not reported with the implication that the examiner also performed the reference standard (surgery) in each of these studies (S. Kim et al., 1999; S. H. Kim et al., 2001; O'Brien et al., 1998). This is suggestive of several sources of bias in these studies, and independently verified results in blinded studies have not since been reported.

From the evidence available, there appears to be some evidence of the diagnostic potential of physical examination tests for identifying SLAP lesions of the shoulder in specific populations, however their diagnostic utility in the primary care setting remains unknown.

Diagnostic Accuracy of Combinations of Clinical Examination Tests.

Several studies investigated the diagnostic accuracy of combinations of physical examination tests for various shoulder conditions, including subacromial pathology (Chew et al.; Litaker et al., 2000; Murrell & Walton, 2001; Park et al., 2005), acromioclavicular joint pathology (Chronopoulos et al., 2004) and glenoid labrum tears (Guanche & Jones, 2003; Joo, Jae, Woo, Hyun, & Ji, 2008; B. W Kibler, Sciascia, Hester, Dome, & Jacobs, 2009; Liu et al., 1996; Oh, Kim, Kim, Gong, & Lee, 2008; Walsworth et al., 2008). In general, combinations of physical examination tests were reported to demonstrate higher levels of accuracy for subacromial, ACJ and GHJ pathology compared with individual tests. Within each of these pathological categories, different combinations of tests were used in each study preventing comparison of the diagnostic accuracy of specific test combinations. An additional finding was the

different methods used for selection of tests to include in test combinations for which diagnostic accuracy was calculated. Main findings from studies investigating the diagnostic accuracy of combinations of clinical examination tests are presented below, with an addition section reporting the methods of selection of clinical variables for inclusion in test combination.

Subacromial pathology

Several studies reported the diagnostic ability of combinations of clinical examination findings for detecting subacromial impingement (Chew et al., 2004; Park et al., 2005) or a rotator cuff tear (Litaker et al., 2000; Murrell & Walton, 2001; Park et al., 2005). In general high levels of accuracy were reported for both these pathology categories using combinations of several clinical tests. Although there were some differences in the tests included in combinations reported to be highly accurate for detecting these pathologies, age, infraspinatus weakness and a painful arc of abduction were common to the majority of these test combinations.

For supraspinatus pathology, age above 39 years, a painful arc during abduction and the report of clicking or popping in the shoulder demonstrated a positive likelihood ratio of 32.2 when all three were present (Chew et al., 2004). The Hawkins-Kennedy, painful arc of abduction and positive infraspinatus muscle test demonstrated a post-test probability of 95% for any degree of subacromial impingement when all three were positive (Park et al., 2005).For a rotator cuff tear, age 65 years or older, night pain and external rotation weakness were more accurate than 'expert diagnosis', and equivalent in accuracy to an MRI (Litaker et al., 2000). Supraspinatus weakness, external rotation weakness and a positive impingement test demonstrated 98% specificity for a rotator cuff tear when all three were present, and the same probability of a rotator cuff tear resulted when age was greater than 60 years and any two of the tests were positive, and a combined absence of these features demonstrated 100% sensitivity for a rotator cuff tear (Murrell & Walton, 2001). A combination of a painful arc during abduction, drop-arm sign and positive infraspinatus muscle test produced a 91% post-test probability for a rotator cuff tear (Park et al., 2005). Although consistently high diagnostic accuracy was reported for a common group of clinical features, all these studies were conducted in samples of patients referred to secondary care services, or who had failed conservative management meaning results are not able to be applied to primary care populations.

Acromioclavicular joint pathology

Only one study was identified in which the diagnostic accuracy of combinations of clinical examination findings was reported for identifying a painful ACJ (Chronopoulos et al., 2004). A painful ACJ in this study was diagnosed when pain was located to the region of the ACJ, tenderness was elicited with palpation of the ACJ and the patient had reported 'complete' or 'near complete' relief of pain following a diagnostic injection of local anaesthetic (Chronopoulos et al., 2004). This study found that combinations of cross-body adduction stress test, the ACJ resisted extension test and the active compression test were 81% sensitive for ACJ pain when at least two of these tests were not positive, and 91% specific for a painful ACJ when all three were positive. However the reference standard consisted of multiple criteria, each with limitations including the lack of information regarding the use of imaging guidance during administration of local anaesthetic into the ACJ, and the uncertain relationship between pain location and localised tenderness, and the response to diagnostic injection.

Glenohumeral joint pathology

Studies reporting diagnostic accuracy of combinations of clinical features for detecting the presence of a glenoid labrum tear reported mixed results. Moderate to high levels of both sensitivity and specificity were reported in three studies for differing combinations of clinical examination tests (Joo et al., 2008; Liu et al., 1996; Walsworth et al., 2008). For the presence of at least one of five tests (apprehension test, relocation test, load and shift test, inferior sulcus sign and crank test) 90% sensitivity and 85% specificity were reported, which was shown to be superior to MRI (59% sensitivity, 85% specificity) for the diagnosis of glenoid labrum tears in this study (Liu et al., 1996). The combination of 'popping' or 'catching', and a positive crank test or anterior slide test demonstrated sensitivity of between 0.82 and 0.89, and specificity between 0.91 and 1.00 for a surgically identified glenoid labrum tear however these results were based only upon those participants from the surgical waiting list who eventually received surgical intervention (Walsworth et al., 2008). When two of three relatively sensitive individual tests (O'Brien, apprehension, or compression-rotation test) were combined with one of three relatively specific tests (Speed's, Yergason's, or biceps load II test), sensitivity and specificity reportedly reached approximately 70% and 95%, respectively (Joo et al., 2008).

In contrast, others found little or no improvement in diagnostic accuracy of combinations of a positive O'Brien test, Jobe relocation test and apprehension test

(sensitivity 0.72, specificity 0.73) compared with individual tests (sensitivity 0.44 to 0.63, specificity 0.73 to 0.87) for detection of a superior labrum anterior-posterior (SLAP) lesion (Guanche & Jones, 2003). Differences in results could be explained by differences in patient populations, test procedures, and also in reference standard procedures (arthroscopic versus open surgical observation). However, the ability of test combinations to accurately identify specific disorders is likely to depend upon the relative association between the individual tests and the 'disease' outcome, and many studies used different methods of selecting the tests that were included in combinations for which diagnostic accuracy was estimated.

Methods used to select clinical test combinations

A number of methods were used to select individual tests for inclusion in test combinations in the diagnostic accuracy studies. Some used statistical techniques such as logistic regression modelling to identify the strongest clinical predictors of specific pathology identified on imaging or surgery (Chew et al.; B. W Kibler et al., 2009; Litaker et al., 2000; Park et al., 2005). Others identified cut-points for the optimal number of tests using area under the receiver operator curve (ROC) (Michener, Walsworth, Doukas, & Murphy, 2009), or simply combined individual tests with the highest sensitivities and specificities (Joo et al., 2008; Oh et al., 2008). The accuracy of test combinations was assessed using criteria in which any one of the tests were positive (Guanche & Jones, 2003; Joo et al., 2008; Liu et al., 1996), absolute number of positive tests were identified (e.g. two of three) (Chew et al.; Murrell & Walton, 2001; Park et al., 2005), minimum numbers of tests were positive (1 or more, 2 or more etc.) (Chronopoulos et al., 2004), all tests, or specific combinations of tests were positive (Joo et al., 2008; B. W Kibler et al., 2009; Michener et al., 2009; Oh et al., 2008; Park et al., 2005; Walsworth et al., 2008). Litaker et al. (2000) assigned points for each positive test based on regression coefficients and assessed the most discriminatory score for rotator cuff tear using ROC analysis (Litaker et al., 2000). No justification was provided for test selection methods or diagnostic accuracy assessment methods, and no studies were identified in which these methods had been prospectively evaluated or compared.

In general, most studies reported improvements in diagnostic accuracy using combinations of clinical tests compared to individual tests (Chew et al.; Murrell & Walton, 2001; Park et al., 2005), however the improvements varied according to the criteria used for assessing diagnostic accuracy. Increases in diagnostic sensitivity were the most commonly reported result of using criteria involving either a single positive

test, or minimum numbers of positive tests for rotator cuff pathology (Michener et al., 2009), painful ACJ conditions (Chronopoulos et al., 2004) and glenoid labrum tears (Walsworth et al., 2008). Sensitivity for detection of a glenoid labrum tear improved when one of a combination of three tests was required to be positive (0.51 to 0.95) compared with individual tests alone (0.43 to 0.50) (Walsworth et al., 2008). However specificity of combinations of tests was reported to be adversely affected using these criteria (Guanche & Jones, 2003; Michener et al., 2009; Walsworth et al., 2008).

Specificity, post-test probabilities and +LR of test combinations was, however, shown to improve when the criterion required higher numbers of positive tests, or all tests to be positive (Chronopoulos et al., 2004; Joo et al., 2008; Park et al., 2005). Specificity of three tests for a full thickness supraspinatus tear improved from 0.70 (one of three tests positive) to 0.98 when all three tests were positive, with +LR increasing from 0.79 to 16.35, significantly increasing the post-test probability for a full thickness tear (Park et al., 2005).

Systematic Reviews of Diagnostic Accuracy Studies

Eight general and narrative reviews were identified (Andrews, 2005; Barber, Field, & Ryu, 2007; Burbank, Stevenson, Czarnecki, & Dorfman, 2008; Pyne, 2004; SooHoo & Rosen, 1996; Tennent, Beach, & Meyers, 2003; Turner-Stokes, 1996; Wilk et al., 2005), as well as 11 systematic reviews that included many of the individual studies identified above (Table 2.3). The majority of the high quality systematic reviews were published immediately prior to commencing this study (2007-2009). It was deemed unnecessary to repeat full systematic reviews for individual diagnostic accuracy studies where these had already been completed to a high standard. In these circumstances, the systematic review search methodology was repeated where possible to update findings, and assess whether any new information was available upon which conclusions of these reviews may be influenced. Two of the systematic reviews also conducted meta-analyses to estimate the pooled accuracy of individual physical examination tests for specific shoulder pathology (Dinnes, Loveman, McIntyre, & Waugh, 2003; Hegedus et al., 2008). A summary of findings from systematic reviews, and findings from new evidence identified from repeating search methods is presented in the following sections relating to subacromial, ACJ and GHJ pathology.

Table 2.3. *Summary of Systematic Reviews of Diagnostic Accuracy Studies*

Study	Design	Quality assessment	Tests evaluated	Target condition
Beaudreuil et al., 2009	SR	None used	physical examination tests	impingement and rotator cuff disease
Meserve et al., 2009	SR/MA	Adapted from Cochrane Methods Group on Systematic Review of Screening and Diagnostic Tests; Irwig et al. (1995).	physical examination tests	SLAP lesions
Munro et al., 2009	SR	QUADAS	physical examination tests	labral pathology
Hughes et al., 2008	SR	Modified NHMRC guidelines	physical examination tests	rotator cuff pathology
Hegedus et al., 2008	SR/MA	QUADAS	physical examination tests	impingement, rotator cuff integrity, glenoid labrum integrity, LHB pathology, ACJ pathology, instability.
Powell et al., 2008	SR	QUADAS	physical examination tests	SLAP lesions
Dessaur and Magarey, 2008	SR	QUADAS	physical examination tests	SLAP lesions
Jones et al., 2007	SR	Not reported	physical examination tests	superior labral lesions
Diehr et al., 2006	G/MA	None used	physical examination tests	rotator cuff tear
Mirkovic et al., 2005	SR	Adapted PEDro scale	physical examination tests	SLAP lesions
Luime et al., 2004	SR	QUADAS	history and physical examination tests	instability and labral lesions
Dinnes et al., 2003	SR	Modified QUADAS	history and physical examination tests	clinical impingement syndrome and rotator cuff tear.

Abbreviations. SR, systematic review; G, general review; MA, meta-analysis; SLAP, superior labrum anterior-posterior; QUADAS, Quality Assessment of Diagnostic Accuracy Studies; NHMRC, National Health and Medical Research Council.

Quality assessment of studies included in systematic reviews

Results of systematic reviews and meta-analyses have important implications for both clinical practice and health policy decision-making, thus the quality of studies included in such analyses is of considerable importance with respect to interpretation of these results. Quality assessment of the studies included in systematic reviews is now recognized as an important aspect of systematic review design (Moher, Liberati, Tetzlaff, & Altman, 2009) and quality assessment was conducted using evidence-based

quality assessment tools in six of the reviews identified (Dessaur & Magarey, 2008; Dinnes et al., 2003; Hegedus et al., 2008; Luime, Verhagen, et al., 2004; Munro & Healy, 2009; Powell, Huijbregts, & Jensen, 2008) (Table 2.3). The review conducted by Hegedus et al., (2008) included assessment of the quality of these studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool (Whiting, Rutjes, Reitsma, Bossuyt, & Kleijnen, 2003) in which a score of 10 or higher was used to represent a high-quality study (Hegedus et al., 2008). The method of quality scoring of studies has been criticised by some citing differences in relative importance of individual items included in scoring tools that are assigned equal value (Whiting, Harbord, & Kleijnen, 2005). It is suggested this may result in incorrect classification of a study as 'high' or 'low' quality when quality cut-points are used (Whiting, Harbord, et al., 2005). However this does provide an indication of overall adherence to study design guidelines, and provides scope for re-evaluation at a later date pending development of validated scoring criteria. The systematic reviews that included quality assessment of diagnostic accuracy studies were prioritised for review, and a summary of their findings relating to specific pathology is presented below.

Subacromial pathology

Several systematic reviews included studies specifically investigating the diagnostic accuracy of physical examination tests for rotator cuff pathology and subacromial impingement (Stage I-III) (Beaudreuil et al., 2009; Diehr, Ison, & Jamieson, 2006; Dinnes et al., 2003; Hegedus et al., 2008; Hughes, Taylor, & Green, 2008). Two reviews assessed the quality of the studies (Dinnes et al., 2003; Hegedus et al., 2008), one conducted a meta-analyses of results for the Neer and Hawkins-Kennedy tests (Hegedus et al., 2008), and another performed a meta-analysis for combinations of impingement tests (Dinnes et al., 2003).

Despite a number of individual studies reporting high sensitivity of the Neer and Hawkins-Kennedy impingement tests (Calis et al., 2000; Leroux et al., 1995; MacDonald et al., 2000), results of the meta-analysis suggested that both the Neer (pooled sensitivity 0.79; pooled specificity 0.53; AUC 0.74) and the Hawkins-Kennedy test (pooled sensitivity 0.79; pooled specificity 0.59; AUC 0.76) were of limited diagnostic value for impingement (Hegedus et al., 2008). In contrast, combinations of impingement tests appeared to produce higher pooled sensitivity (0.91) for impingement pathology (Dinnes et al., 2003). Several tests were identified with more than 80% specificity for impingement pathology (drop arm test, Yergason's test, Speed's test and

passive external rotation test), however many studies were rated only moderate to low quality predominantly including low sample sizes, non-consecutive samples and almost none reported handling of missing or indeterminate results (Dinnes et al., 2003). Conclusions from systematic reviews were consistent, and agreed that based upon the results of higher quality diagnostic studies, the Hawkins-Kennedy test or combinations of impingement tests may serve as a screen, and the empty can or infraspinatus test may help to confirm the diagnosis of stage I to stage III impingement (Dinnes et al., 2003; Hegedus et al., 2008).

For assessing integrity of the rotator cuff, only half of the studies in the Hegedus et al. (2008) review were considered of 'high quality', and none of the tests investigated in these studies were shown to be consistently diagnostic for a tear of the rotator cuff. However the external rotation lag sign and the drop arm tests demonstrated high specificity in several studies for a tear of any rotator cuff component (0.72 to 0.98) (Hegedus et al., 2008). The belly press test was also identified by two systematic reviews to be of value for ruling-in a subscapularis muscle tear (Hegedus et al., 2008; Hughes et al., 2008).

One new study was identified in which limited diagnostic value for the shoulder shrug test was reported for a rotator cuff tear (Jia, Ji, Petersen, Keefer, & McFarland, 2008). Two new studies investigated previously reported lag signs for the diagnosis of a rotator cuff tear and results supported previous findings of high specificity (98%) of the external rotation lag sign (Castoldi et al., 2009), and the other reporting 100% sensitivity of the internal rotation lag sign for full thickness rotator cuff tears. However again, methodological limitations including small sample sizes (;Miller, 2008 #568}, retrospective design (Castoldi et al., 2009; Jia et al., 2008), use of different reference standard procedures within the same sample (Castoldi et al., 2009) and lack of reported confidence intervals for diagnostic estimates (Miller et al., 2008a) mean results are not likely to alter conclusions drawn from existing systematic reviews.

There appears to be general agreement that for assessing integrity of the rotator cuff, a positive external rotation lag sign may be of diagnostic value for any rotator cuff tear, or an infraspinatus tear and the belly press test may be of value for ruling-in a subscapularis tear (Hegedus et al., 2008). Both systematic reviews reported that convincing evidence for the diagnostic accuracy of individual tests for rotator cuff pathology was lacking due to the poor quality of studies included in these reviews.

Acromioclavicular joint pathology.

Only one systematic review included studies investigating the accuracy of physical examination tests for ACJ pain or pathology (Hegedus et al., 2008). This systematic review identified only two studies of high design quality, and identified a trend of decreasing diagnostic accuracy results with increasing study design quality. However, they suggested that the absence of tenderness to palpation of the ACJ may help to rule-out ACJ pathology and a positive active compression test may help to rule-in AC joint pathology (Hegedus et al., 2008). No additional diagnostic accuracy studies were identified in which ACJ pain or pathology was used as the outcome condition of interest.

Glenohumeral joint pathology.

Eight systematic reviews were identified that were published between 2007 and 2009, that included studies investigating the diagnostic accuracy of physical examination tests for GHJ pathology, with all reviews focusing on glenoid labral lesions (Dessaur & Magarey, 2008; Hegedus et al., 2008; Jones & Galluch, 2007; Luime, Verhagen, et al., 2004; Meserve, Cleland, & Boucher, 2009; Munro & Healy, 2009; Powell et al., 2008). Quality assessment using the QUADAS instrument was used in all but three of these reviews (Jones & Galluch, 2007; Meserve et al., 2009; Mirkovic, Green, Taylor, & Perrott, 2005). No systematic reviews were identified in which studies investigated clinical examination test accuracy for other GHJ pathologies.

Common findings from the systematic reviews included the potential diagnostic value of the biceps load II test (Dessaur & Magarey, 2008; Hegedus et al., 2008; Jones & Galluch, 2007; Meserve et al., 2009; Munro & Healy, 2009) and the Speed's/palm-up test (Meserve et al., 2009; Powell et al., 2008) to confirm the presence of a SLAP lesion in specific populations. Others supported the use of the O'Brien's test to rule-out a SLAP lesion (Meserve et al., 2009) anterior apprehension and Jobe relocation tests to rule-in a SLAP lesion (Powell et al., 2008)and the Kim and Jerk test to differentiate labral pathology from other pathologies in select populations (Munro & Healy, 2009).

Several new studies were also identified in which the accuracy of new physical examination tests was reported for labral lesions and instability (Bushnell, Creighton, & Herring, 2008; B. W Kibler et al., 2009; Y. S. Kim et al., 2007; Schlechter, Summa, & Rubin, 2009). Of the new studies identified using previous search methods(Hegedus et al., 2008), methodological limitations were present, including small sample size (n=29) (Bushnell et al., 2008), only one third received the reference standard procedure (B. W

Kibler et al., 2009), and retrospective analysis of arthroscopy findings and inadequately described blinding procedures (Schlechter et al., 2009). The highest quality study reported moderately high levels of reproducibility (kappa 0.77), sensitivity (0.82) and specificity (0.86) of the passive compression test for the diagnosis of SLAP lesions although confidence intervals were not reported (Y. S. Kim et al., 2007). One new study was identified in which sensitivity analysis was performed on those with arthroscopically proven SLAP lesions using previously reported tests for SLAP lesions (Pandya, Colton, Webner, Sennett, & Huffman, 2008). Results indicated 90% sensitivity for the active compression test, 80% sensitivity for the dynamic shear test, 76% sensitivity for the Jobe relocation test and 100% sensitivity when any one of these tests was positive, however specificity could not be estimated using this methodology and patients with co-existing pathology identified during arthroscopy were excluded limiting the application of results in clinical practice.

Results of the new studies identified do not alter conclusions drawn in systematic reviews, with the possible exception of the passive compression test providing a potentially useful diagnostic test for SLAP lesions in orthopaedic settings. Combined evidence from systematic reviews and additional studies suggests that the biceps load II test and the Speed's/palm up test may be of diagnostic value for identifying SLAP lesions, and the Kim and Jerk tests may be of diagnostic value for other lesions of the glenoid labrum in secondary care or orthopaedic settings.

Methodological Quality of Diagnostic Accuracy Studies

A diagnostic test can be considered useful when it helps differentiate conditions that may prompt clinical actions such as further diagnostic testing, or the initiation, modification, or termination of treatment (Bossuyt et al., 2003). Several potential sources of bias and variation respectively affect the internal and external validity of diagnostic accuracy studies which could lead to inappropriate acceptance of a test into clinical practice resulting in an incorrect diagnosis and treatment pathway (Whiting et al., 2004). Despite the availability of several tools that provide evidence-based guidelines for the design (QUADAS (Quality Assessment of Diagnostic Accuracy Studies), reporting and interpretation of diagnostic accuracy studies (STARD statement (Standards for Reporting of Diagnostic Accuracy)), (Bossuyt et al., 2003; Whiting et al., 2003), methodological quality of diagnostic studies remains generally poor, or inadequately described (Reid, Lachs, & Feinstein, 1995; Rutjes et al., 2006; Sheps & Schechter, 1984; Whiting, Rutjes, et al., 2005). Findings of the literature review

performed as part of this thesis support these findings, with several common methodological limitations identified in the majority of studies that may affect both the diagnostic accuracy estimates, as well as the applicability of results in clinical practice.

Internal and external validity of diagnostic studies.

Bias or variation may be present in any aspect of the study including quality of study design, conduct, analysis and reporting. Differences in patient selection methods, the test protocols (both index tests and reference standard tests), the process of verification using the reference standard and the way the index test and reference standard are reviewed may all introduce bias into the study, resulting in estimates of diagnostic accuracy of the test that may differ from the true performance of the test (Whiting et al., 2004). The presence of bias may specifically affect test sensitivity resulting in the potential for overestimation of sensitivity and overall diagnostic accuracy (Lijmer et al., 1999; Whiting et al., 2004). Several aspects of study design may also limit the population to which results may be applied (variability), and sources of variation include differences between study populations and settings, test protocols or criteria used to define the target disorder (Whiting et al., 2004). Many of the diagnostic accuracy studies identified in the literature search contained various sources of bias and variation that may affect interpretation of reported test accuracy, or application of results.

Design and sampling

Many studies utilised retrospective design recruiting a non-consecutive sample of patients who had already undergone the surgical reference standard procedure (selection bias). Many studies also involved only small or moderate sample sizes (less than 100 participants), and reporting of confidence intervals that provide an indication of the precision of diagnostic accuracy estimates was lacking in many studies.

Disease spectrum bias represents another threat to external validity of many studies where inclusion criteria were patients either currently on surgical waiting lists, or those who had already undergone surgical procedures. It is likely that the necessity for surgery represents the severe end of the disease spectrum, and results cannot therefore be generalised to those other populations with wider variability in disease status, including minor shoulder complaints.

Reference standard tests

Surgical confirmation of pathology was used as the reference standard in the majority of diagnostic studies, with a small number using either a diagnostic injection

(Calis et al., 2000; Chronopoulos et al., 2004; Walton et al., 2004) or imaging investigations including diagnostic ultrasound (Miller et al., 2008b), MRI or MRA as the reference standard (Ardic et al., 2006; Calis et al., 2000; Itoi, Kido, Sano, Urayama, & Sato, 1999; Mimori et al., 1999; O'Brien et al., 1998; Scheibel, Magosch, Pritsch, Lichtenberg, & Habermeyer, 2005). The use of imaging or surgery as the reference standard in these studies, while permitting visualisation of pathological changes, does not take into account to what extent the observed pathology contributed to symptoms. Although considered the 'best available' reference standard tests for identification of pathological changes, surgery and imaging findings may lack an element of face validity for identification of symptomatic lesions. The use of local anaesthetic injections when performed in the absence of imaging guidance (Calis et al., 2000)has also been criticised due to the inability to confirm accurate location of the targeted structure and lack of specificity resulting from the potential anaesthetisation of structures other than the intended site (Hughes et al., 2008).

In many studies, the reference standard was inadequately described. All studies were conducted in the orthopaedic setting, and in many cases the orthopaedic surgeon performing the reference standard was a highly experienced specialist in shoulder surgery. Differences in the prevalence of specific conditions, disease severity, examiner training and experience mean results of these studies cannot be generalised to the primary care setting, or within the orthopaedic setting where less experienced surgeons perform the clinical test procedures.

Interpretation

Reviewer bias could not be ruled out in a number of studies, in which it was either not stated, or it was unclear whether the examiner who performed the index test was the same examiner who performed the reference standard procedure (surgery in most cases). The majority of studies did not report the reliability of their test procedures, and only one study reported internal (within-study) validation of their findings (Litaker et al., 2000). No studies reported prospective validation of their results.

Analysis

Almost without exception, the frequency and handling of uninterpretable or indeterminate results was not reported. Exclusion of cases from the analysis in which either the index test result was unclear or the test could not be performed, or results of reference standard procedure were indeterminate may lead to biased estimates of sensitivity or specificity.

Key Findings

- There is a lack of studies reporting the diagnostic accuracy of clinical examination tests for specific shoulder pathology in primary health care populations and the performance these tests in this setting remains unknown.

- Evidence-based tools are available to guide the design, conduct and reporting of high quality diagnostic accuracy studies but despite this the methodological design and/or reporting of diagnostic accuracy studies remains generally poor.

- The majority of studies investigated orthopaedic tests, with few studies investigating the diagnostic accuracy of patient history or other aspects of physical examination including range of motion and resisted tests and from high quality studies, at best there is limited evidence for the diagnostic value of only a few physical examination tests for specific shoulder conditions.

- There is some evidence that combinations of clinical examination findings may improve diagnostic accuracy compared with individual tests.

- The reference standards used in diagnostic studies were primarily high-tech imaging (MRI) or surgery, which do not take into account the contribution of observed pathology to symptoms.

2.4. Diagnostic Imaging

In this section, literature was reviewed to identify evidence for the relevance of diagnostic imaging findings to the diagnosis of symptomatic subacromial, ACJ and GHJ conditions.

Search Terms

Literature was searched using the concepts and search terms listed in Table 2.4.

Literature Search Results

A limited number of studies (25) were identified in which the relationship between imaging findings and symptoms had been reported in both symptomatic and asymptomatic patient groups for subacromial, ACJ and GHJ pathology. However, of the studies involving symptomatic participants, none evaluated the diagnostic accuracy of the imaging findings for identifying the presence or degree of symptoms using pain-relieving procedures such as diagnostic injections of local anaesthetic. Key findings from the studies reviewed are presented below.

Table 2.4. *Search Terms for Diagnostic Imaging*

Concept	Combine search terms	Subject headings and search terms	Result
Shoulder pain		See "shoulder pain" (Table 2.1)	47645
Diagnostic imaging	OR	diagnostic imaging.sh exp ultrasonography/ ultraso*.mp OR songra*.mp OR US.mp OR "diagnostic ultraso*".mp exp radiography/ exp magnetic resonance imaging "magnetic resonance imaging".mp OR "MR imaging".mp OR MRI.mp "magnetic resonance arthrogr*".mp OR "MR arthrogr*".mp OR MRA	6308093
Diagnostic value	OR	sensitiv*.mp. diagnos*.mp. di.fs. valid*.tw OR accura*.tw OR specific*.tw OR "likelihood ratio".tw reliab*.mp OR varia*.mp OR agreement.mp diagnos* ADJ5 valu*.tw symptom* OR asymptomatic	5393923
Concepts	AND	Diagnostic imaging Diagnostic value	142
		Limits (English language, humans, year 1940-2009)	97
		Abstract and title search	25

Relationship between Imaging Findings and Symptoms

Of the limited number of studies available in which the symptomatic relevance of imaged pathology was investigated, the majority related to the prevalence of rotator cuff tears and ACJ pathology in asymptomatic populations, and the presence of glenoid labrum tears in specific athletic populations. There appears to be a lack of studies in which the relevance of imaged pathology in symptomatic primary care populations was investigated.

Subacromial pathology

Age is reported to be associated with the presence of rotator cuff tears in asymptomatic populations. The prevalence of asymptomatic rotator cuff tears in the general population is reported to be between 23% and 34% (Sher, Uribe, Posada, Murphy, & Zlatkin, 1995; Tempelhof, Rupp, & Seil, 1999). Asymptomatic rotator cuff tears are reported to become increasingly prevalent with advancing age. The prevalence of rotator cuff tears was shown to increase from 20% to 50% in those over 60 years of age (Milgrom et al., 1995; Sher et al., 1995), and to between 51% and 80% in those over 80 years of age (Milgrom et al., 1995; Tempelhof et al., 1999). No full thickness rotator

cuff tears were identified in a sample of asymptomatic 19 to 39 year old participants, and only one partial thickness tear was identified in this group (Sher et al., 1995). No studies were found in which the prevalence of rotator cuff tears in a cohort suffering a current episode of shoulder pain was reported.

It is unclear from the evidence reviewed which imaging features distinguish symptomatic from asymptomatic rotator cuff tears. Increasing tear size was related to an increase in superior humeral head migration seen in rotator cuff deficient shoulders, and greater amounts of superior humeral head migration were reported in symptomatic compared with asymptomatic tears (Keener, Wei, Kim, Steger-May, & Yamaguchi, 2009). Larger bursal effusions on MRI were reported in symptomatic compared with asymptomatic rotator cuff tears, although this study used a small sample size and statistical significance was not achieved (Hirano, Sashi, Izumi, Itoi, & Watarai, 2006). A subacromial bursal effusion/hypertrophy on MRI was also associated with higher levels of disability in a study involving patients with impingement syndrome (r=0.5, p=0.03) (Ardic et al., 2006). However the majority of these studies included patients from orthopaedic care settings, and the relationship between imaging findings and symptoms of less severe shoulder conditions was lacking.

Approximately 51% of asymptomatic patients with known rotator cuff tears were reported to develop symptoms within a 3 year period, and progression of tear size was postulated as a possible cause for the development of symptoms from previously asymptomatic patients (Yamaguchi et al., 2001). In contrast, no statistical relationship was found between the level of pain and disability and the size and location of full-thickness tears of the rotator cuff on MRI (Krief & Huguet, 2006).

Acromioclavicular joint pathology.

Asymptomatic degenerative changes affecting the ACJ are reported to be common. Among asymptomatic community volunteers in a 'sports medicine' study, ACJ osteoarthrosis was present in 76% of shoulders (Needell & Zlatkin, 1997). Shubin-Stein (2001) report a similarly high prevalence of degenerative ACJ changes in an asymptomatic cohort under the age of 30 years (68%) with a substantial increase in degenerative changes (93%) in those over 30 years of age (Shubin Stein et al., 2001). When compared with an asymptomatic sample, degenerative changes were reported to be more advanced in the symptomatic group, and MRI evidence of bone oedema in the lateral clavicle or acromion (or both) were only present in symptomatic participants (Shubin Stein, Ahmad, Pfaff, Bigliani, & Levine, 2006). The finding of bone oedema in

the clavicle or acromion was therefore reported to be a more reliable indicator of symptomatic ACJ pathology than degenerative changes on MRI (Shubin Stein et al., 2006).

Glenohumeral joint pathology.

The majority of studies investigating GHJ pathology were conducted either in orthopaedic settings or in specific athletic populations comparing symptomatic with asymptomatic subgroups (Funk & Snow, 2007; Haddock & Funk, 2006; McFarland, Tanaka, Garzon-Muvdi, Jia, & Petersen, 2009; Miniaci, Mascia, Salonen, & Becker, 2002; Oh et al., 2008). A minimal biceps tendon sheath effusion on MRI, reported to be indicative of a glenohumeral joint effusion, was reported to be present in 79% of asymptomatic volunteers (age 19 to 88 years) (Needell & Zlatkin, 1997). Asymptomatic glenoid labrum abnormalities were also reported to be present on MRI in 79% of professional baseball pitchers (Miniaci et al., 2002). In symptomatic shoulders who were awaiting physiotherapy for suspected impingement syndrome, more painful shoulders (r=0.8, *p*=0.000) and higher levels of disability measured on the Disability of Arm, Shoulder and Hand questionnaire (r=0.6, *p*=0.02) had more frequent glenoid labrum tears on MRI (Ardic et al., 2006).

Results of these studies suggest that glenoid labrum lesions may not always be related to symptoms of should pain in throwing populations, but labral tears may be of more significance in those with suspected impingement. However the small number of available studies and the lack of imaging studies in symptomatic populations mean firm conclusions cannot be drawn regarding the relevance of imaged GHJ pathology to symptoms in this population.

Overall, these studies indicate that in the asymptomatic population, there is a high prevalence of rotator cuff and ACJ pathology which increases with age, and in throwing populations, asymptomatic glenoid labrum lesions are common. While studies involving asymptomatic participants provide information regarding the prior probability of imaging findings in specific populations, information regarding imaging findings that are associated with painful shoulder conditions seen in primary care is scarce.

Key Findings

- There is an age-related increase in prevalence of asymptomatic rotator cuff tears and degenerative ACJ changes.

- Little evidence exists reporting the relationship between diagnostic imaging findings of subacromial, ACJ and GHJ pathology, and symptoms of shoulder pain.
- No estimates of diagnostic accuracy were found for the ability of imaging findings to predict the presence of symptomatic pathology.

2.5. Reference Standards for Identifying Sources of Shoulder Pain

The previous literature review sections identified a large number of studies in which radiological imaging or surgery was used to identify pathology. However the contribution of observed pathology to symptoms cannot be evaluated using these methods. Pain is a physiologic phenomenon that cannot be visualised on imaging, hence a physiologic test such as a diagnostic block using local anaesthetic is required to identify its presence (Bogduk, 2009). Available literature was reviewed to evaluate the use of diagnostic blocks of local anaesthetic for identifying the tissue source of pain and to summarise issues relating to the validity of these procedures for use in diagnostic research.

Search Terms

The following terms (Table 2.5) were used to identify relevant literature relevant to musculoskeletal source of pain.

Table 2.5. *Search Terms for Reference Standards for Identifying Sources of Shoulder Pain*

Concept	Subject headings and search terms	Result
Diagnostic injections	anesthetics, local.sh AND du.fs	163
	Limits (English language, humans, year 1940-2009)	51
	Abstract and title search	32

Literature Search Results

The literature search revealed very few studies in which the use of diagnostic injections for identifying sources of shoulder pain was described (Calis et al., 2000; Chronopoulos et al., 2004; Strobel, Pfirrmann, Zanetti, Nagy, & Hodler, 2003; Walton et al., 2004). The vast majority of literature, including four systematic reviews, relates to the use, and validity of diagnostic injections for painful spinal, hip and ankle joint

conditions (Datta, Lee, Falco, Bryce, & Hayek, 2009; Falco et al., 2009; Manchikanti, Dunbar, et al., 2009; Manchikanti, Glaser, Wolfer, Derby, & Cohen, 2009). Practice guidelines were also identified published by the International Spinal Intervention Society (ISIS) (2004), in which evidence-based summaries of relevant literature were presented as accepted practice guidelines for diagnostic and therapeutic spinal interventions (International Spine Intervention Society [ISIS], 2004).

Key findings from those studies identified relating to the use of diagnostic injections for shoulder pain are presented, and other relevant methodological issues from other studies relating to the validity of these procedures are also summarised.

Use of Diagnostic Blocks for Shoulder Pain.

Subacromial structures, the ACJ and capsular and intra-articular GHJ pathology have been confirmed as sources of shoulder pain using injections of saline (Gerber, Galantay, & Hersche, 1998; Larson, O'Connor, & Nirschl, 1996). Diagnostic injections have also been advocated for the clinical diagnosis of 'subacromial impingement' (Calis et al., 2000; Cyriax, 1978; Neer, 1972) and to assist in differentiating local shoulder pain from spinal or visceral sources of referred pain (Larson et al., 1996), as well as to help identify those patients who may benefit from CSI (Neer, 1972). Several studies were identified in which subacromial diagnostic injections had used as a reference standard procedure in studies investigating the diagnosis of subacromial impingement syndrome (Calis et al., 2000), and painful ACJ conditions (Chronopoulos et al., 2004; Strobel et al., 2003; Walton et al., 2004).

Reports of the use of glenohumeral joint diagnostic blocks are less common, however therapeutic interventions are frequently performed as part of the MRA procedure for adhesive capsulitis that include introduction of local anaesthetic including distension arthrography and intra-articular corticosteroid injection (Andren & Lundberg, 1965; Mulcahy, Baxter, Oni, & Finlay, 1994). No studies were identified in which response to an intra-articular injection of local anaesthetic into the GHJ was used as a diagnostic test. Although its use as a diagnostic test using a post-injection pain response has not been described, studies suggest the increased volume of anaesthetic used by some practitioners is indicative of its use for diagnostic purposes where reduced joint capacity may be indicative of adhesive capsulitis (Skedros, Hunt, & Pitts, 2007).

Validity of Diagnostic Blocks for Shoulder Pain

Available evidence was consistent in reporting that the use of imaging guidance improved injection accuracy for targeting specific shoulder structures. Studies of subacromial injections in the absence of imaging guidance report successful and accurate infiltration of the subacromial bursa in only 46% to 83% of cases, with other structures, including the rotator cuff and glenohumeral joint, also frequently infiltrated (Henkus, Cobben, Coerkamp, Nelissen, & Van Arkel, 2006; Kuhn & McGuigan, 2006; Partington & Broome, 1998; Yamakado, 2002). Despite the use of intra-articular ACJ injections as the reference standard in diagnostic studies (Chronopoulos et al., 2004; Strobel et al., 2003; Walton et al., 2004), imaging guidance was not used in all cases, and anaesthetic response criteria varied considerably making comparisons difficult.

Accuracy of ACJ injections performed in the absence of imaging guidance is reported to be variable with successful infiltration rates for the ACJ reported to vary between 39% and 67% (Bisbinas, Belthur, Said, Green, & Learmonth, 2006; Partington & Broome, 1998). There is also general agreement in the spine pain literature that the use of contrast-enhanced fluoroscopic guidance to ensure accuracy of needle placement and assess containment of injectate within the targeted structure during intra-articular injections is required to ensure face validity of these procedures (Bogduk, 2004b, 2005; Bogduk, Dreyfuss, & Govind, 2009; Manchikanti, Derby, et al., 2009; Manchikanti, Pampati, & Cash, 2010).

In the spine pain literature, false-positive rates following zygapophyseal joint nerve blocks have been reported using placebo-controlled or double anaesthetic block procedures (Manchikanti, Pampati, Fellows, & Bakhit, 1999, 2000; Schwarzer, Aprill, et al., 1994), however there appears to be a lack of evidence regarding the false-negative or false-positive rates following diagnostic shoulder injections.

Evaluating the Response to Diagnostic Blocks

Following injection of local anaesthetic, the change in pain intensity based upon clinical reassessment of provocative tests is reported to be the accepted criterion for a positive anaesthetic response (International Spine Intervention Society [ISIS], 2004). The use of serial 100mm visual analogue scales (VAS) is reported to be the accepted method of measuring change in pain intensity following diagnostic block procedures (International Spine Intervention Society [ISIS], 2004). Of the studies in which a diagnostic block was used to identify shoulder pain details regarding the criteria used to

define a positive response to diagnostic injection were either inadequate or not reported (Calis et al., 2000; Chronopoulos et al., 2004; Strobel et al., 2003; Walton et al., 2004).

Spinal intervention practice guidelines suggest that diurnal fluctuations in pain intensity may be up to 20mm as measured on the VAS, and that reports of post-injection change in pain intensity within this range may not be attributable to the diagnostic block procedure (International Spine Intervention Society [ISIS], 2004). The implication is that participants with low levels of pain may record changes in pain that are not attributable to the diagnostic block procedure.

Positive anaesthetic response criteria.

Response to diagnostic block is considered positive when there is a reported reduction in pain severity following the injection of local anaesthetic (Bogduk, 2004b). In studies in which diagnostic blocks were used to identify shoulder pain, subjective criteria including "complete or near-complete relief" (Chronopoulos et al., 2004) or "marked" improvement in pain (Calis et al., 2000) were used as the positive anaesthetic response criteria, with a 50% (Walton et al., 2004), and 70% reduction in pain (Strobel et al., 2003) used as the positive response criteria by others. Considerable variation exists in other studies regarding the amount of pain relief considered to represent a positive anaesthetic response. Studies investigating spinal pain have reported a range of cut points including 50% (S. Carette et al., 1991; Walton et al., 2004), 70% (Broadhurst, 1989; Strobel et al., 2003), 75% (Manchikanti et al., 1999; Revel et al., 1998), 80% (Dreyfuss et al., 2000; Laslett, 2006; Laslett, Aprill, McDonald, & Young, 2005), 90% (Dreyfuss, Michaelsen, Pauza, McLarty, & Bogduk, 1996), and proportions of patients satisfying a range of criteria (75% to 95%) have also been reported (Laslett, McDonald, Aprill, Tropp, & Oberg, 2006).

There also appears to be considerable debate regarding the percentage cut point for pain relief following diagnostic injection that represents justification for performance of a therapeutic spinal interventional technique, with higher cut-points assumed to represent higher levels of diagnostic certainty for which the intervention is to be applied. It has been argued that at least 80% pain relief should be reported to justify the cost of performing expensive procedures such as radiofrequency neurotomy (Manchikanti et al., 2010). The use of more stringent criteria such as 80% has been shown to reduce the false-positive response rate in patients with confounding factors reporting lumbar spine pain (Datta et al., 2009). Higher cutpoints have also been shown to produce high specificity with regard to identification of the tissue origin of pain

(Laslett et al., 2005), and 80% pain relief is the criterion used in the majority of higher quality studies selected for systematic review of diagnostic efficacy (Datta et al., 2009; Falco et al., 2009). Greater diagnostic stability for lumbar facet joint pain at 2-year follow-up has also been reported using the 80% pain relief criterion (90% of participants) compared with the lower pain relief criterion of 50% (51% of participants) (Manchikanti et al., 2010). No such studies were identified in which the relationship between anaesthetic response, diagnostic stability or treatment outcome has been investigated for diagnostic injections around the shoulder.

Key Findings

- Injection accuracy for subacromial and ACJ injections appears to be improved with the use of imaging guidance.
- Although the intra-articular injection of anaesthetic into the glenohumeral joint is used in interventional procedures for shoulder pain, its use as a diagnostic test is not well described.
- The use of the 100mm VAS is the accepted method of recording pre- and post-injection pain intensity.
- Patients who report low pre-injection pain intensity (<20mm) may report post-injection change in pain intensity that represents diurnal fluctuation in pain scores and may not be attributable to the anaesthetic.
- Positive anaesthetic response criteria following shoulder injections are widely variable, however it appears that an 80% reduction in pain following injection of local anaesthetic is the most commonly accepted positive anaesthetic response criterion from established procedures.
- The false-positive and false-negative rates of diagnostic blocks around the shoulder have not been reported.

2.6. Reference Standards for Identifying Shoulder Pathology

In clinical practice, methods used to visualise shoulder pathology include plain film x-ray, diagnostic ultrasound, conventional MRI, MRA, computerized tomography and, in some cases, arthroscopic or open surgery are used to assess shoulder pathology.

Diagnostic ultrasound is one of the most commonly used imaging modalities and is widely available in primary health care settings. It is quick, portable, non-invasive and inexpensive, does not involve ionising radiation. Diagnostic ultrasound has the

benefit of enabling high resolution scanning and dynamic assessment of subacromial structures (McNally, 2004). More advanced imaging investigations such as conventional MRI and magnetic resonance arthrogram (MRA) are also becoming increasingly available, providing improved visualisation of pathologies such as glenoid labral lesions and tendon pathology (Shahabpour, Kichouh, Laridon, Gielen, & De Mey, 2008). The diagnostic accuracy of these procedures for identifying specific shoulder pathology was investigated to assess their suitability for use as reference standard procedures for identification of subacromial, ACJ and GHJ pathology, and to evaluate their ability to detect other imaged pathology that may be used as index tests in the diagnostic accuracy study analysis.

Search Terms

A literature search was undertaken using the concepts and search terms in Table 2.6 to evaluate the evidence for diagnostic accuracy of diagnostic ultrasound, MRI and MRA compared with surgical findings that are considered to be the best available reference standard for identification of shoulder pathology (Paavolainen & Ahovuo, 1994).

Literature Search Results

A large number of individual studies were identified in which the accuracy of diagnostic ultrasound, magnetic resonance imaging (MRI) and magnetic resonance arthrogram (MRA) had been compared with surgery for identification of specific shoulder pathology including subacromial, ACJ and GHJ pathology. In addition, twenty five reviews, including two recent systematic reviews were identified (de Jesus et al., 2009; Dinnes et al., 2003). These two high quality systematic reviews included many of the individual studies also identified in the literature search and it was deemed unnecessary to repeat the systematic reviews, however the search strategy was repeated for articles published between 2007 and 2009 to update results of one of the systematic reviews (de Jesus et al., 2009).

Results of individual studies and systematic reviews are presented in sections that summarise the ability of each imaging investigation procedure to detect subacromial, ACJ and GHJ pathology by comparison with a surgical reference standard. Results for the validity of ultrasound and MRI are presented, followed by a section that compares MRI with MRA as many studies directly compared the two procedures in the same cohort.

Table 2.6. *Search Terms for Reference Standards for Identifying Shoulder Pathology*

Concept	Combine search terms	Subject headings and search terms	Results
Shoulder pain		See "shoulder pain" (Table 2.1)	47645
Diagnostic accuracy		See "diagnostic accuracy" (Table 2.2)	4151780
Diagnostic imaging	OR	diagnostic imaging.sh exp ultrasonography/ OR ultrason*.mp exp radiography/ OR rad*.mp exp magnetic resonance imaging/ OR "magnetic resonance imaging".mp OR "MR imaging".mp OR MRI.mp "magnetic resonance arthrogr*".mp OR "MR arthrogr*".mp OR MRA	1000580
Concepts	AND	Shoulder pain Diagnostic accuracy Diagnostic imaging	2018
		Limits (English language, humans, year 1990-2009)	1052
		Title and abstract search	123

Validity of Diagnostic Ultrasound Imaging

The majority of individual studies investigated the reliability and diagnostic accuracy of ultrasound imaging for rotator cuff tears, however a small number of studies were also identified investigating the diagnostic accuracy of ultrasound imaging for detecting tendinopathy (Martin-Hervas, Romero, Navas-Acien, Reboiras, & Munuera, 2001; Naredo et al., 1999), calcific tendon lesions (Kayser, Hampf, Pankow, Seeber, & Heyde, 2005; Martin-Hervas et al., 2001), acromioclavicular joint pathology (Scheel et al., 2005; Schmidt, Schmidt, Schicke, & Gromnica-Ihle, 2004), and glenohumeral joint pathology (Hammar et al., 2001; Lange, Piegsa, Teichmann, & Neeck, 2000; Taljanovic et al., 2000). Summaries of main review findings for each of these shoulder pathologies are presented below.

Advances in diagnostic ultrasound technology and expertise over time mean results of earlier published studies may no longer accurately reflect the ability of diagnostic ultrasound to identify specific pathology (Teefey et al., 2004). For shoulder pathologies in which a large number of articles were identified (rotator cuff tears), studies were prioritised for review in which higher frequency equipment was used (minimum 7.5 to 10 MHz) that are more commonly used in radiology practice due to their superior imaging ability (Teefey et al., 2004).

Subacromial pathology

A large number of studies were identified in which the diagnostic accuracy of ultrasound was investigated for rotator cuff tears, including two systematic reviews (de Jesus et al., 2009; Dinnes et al., 2003). A smaller number of studies were identified investigating the ability of ultrasound to detect other subacromial pathology including subacromial bursitis, dynamic bursal impingement, tendinopathy and calcific lesions.

Subacromial bursitis.

Several studies assessed the ability of ultrasound to diagnose subacromial bursitis compared with MRI or surgery (Awerbuch, 2008; Bruyn et al., 2009; Bruyn et al., 2010; Bureau, Beauchamp, Cardinal, & Brassard, 2006; Dinnes et al., 2003; Farin, Jaroma, Harju, & Soimakallio, 1990; Kayser et al., 2005; King & Healy, 1999; Le Corroller, Cohen, Aswad, Pauly, & Champsaur, 2008; Naredo et al., 1999; Naredo et al., 2006; Scheel et al., 2005; Schmidt et al., 2004). Diagnostic ultrasound imaging is reported to visualise several characteristics of the SAB including bursal fluid or effusion, synovial hypertrophy (thickening), and bursal 'bunching' under the acromion during dynamic assessment (Kolla & Motamedi, 2007). However, diagnostic criteria for subacromial bursitis in these studies were variable, or inadequately reported. Differences most commonly related to the bursal dimension (thickness) that was considered pathological and this varied from 2mm (Kolla & Motamedi, 2007; Naredo et al., 2002; M. Van Holsbeeck & Strouse, 1993) to 3mm thickness (Bruyn et al., 2009).

The lack of accepted diagnostic criteria for subacromial bursitis is one of the likely explanations for the variable reported agreement between musculoskeletal ultrasound experts for identification of SAB pathology on ultrasound (kappa 0.50 to 0.97; agreement 84% to 97%) (Bruyn et al., 2009; Bruyn et al., 2010; Le Corroller et al., 2008; Naredo et al., 2006; Schmidt et al., 2004). Among a group of musculoskeletal ultrasound experts, most disagreements related to variations in dynamic assessment and judgement of SAB fluid as being normal or pathological (Naredo et al., 2006).

The sensitivity of ultrasound for the diagnosis of subacromial bursitis compared with surgical findings was reported to range from 0.71 to 0.79 and specificity from 0.96 to 0.98 (Farin et al., 1990; Kayser et al., 2005). These studies were conducted in different settings (radiology and orthopaedic departments), using different reference standards (arthroscopy and open surgery). Despite the moderate to high levels of diagnostic accuracy reported in these studies, the lack of expert consensus upon the diagnostic criteria for subacromial bursitis, the variable agreement between expert

examiners, and the small number of diagnostic accuracy studies, ultrasound imaging does not appear to be an appropriate reference standard for the diagnosis of subacromial bursitis at this time.

Subacromial impingement.

Validity of dynamic ultrasound assessment of subacromial impingement was assessed in a small number of studies. Diagnostic criteria for dynamic impingement were consistently reported as lateral pooling of fluid within the SAB during abduction, or alteration of the normally convex surface of the subacromial bursa alone or of the subacromial bursa and of the supraspinatus tendon when the greater tuberosity of the humeral head passed underneath the acromion (Bureau et al., 2006; Read & Perko, 1998). Osseous impingement was also described in two studies as superior migration of the head of the humerus preventing passage of the greater tuberosity under the acromion (Bureau et al., 2006; Read & Perko, 1998) and only one study correlated the observed impingement with participant symptom response (Bureau et al., 2006).

Few studies investigated the reliability of examiners in reporting subacromial impingement, however 88% agreement has been reported for this diagnosis among a group of experts in musculoskeletal ultrasound (radiologists and rheumatologists) from different countries (Naredo et al., 2006). In studies investigating the diagnostic accuracy of "impingement" diagnosed during dynamic ultrasound assessment, sensitivities were reported to range between 0.71 and 0.79, and specificities between 0.88 and 0.96 in which surgical confirmation of "impingement" was used as the reference standard (Farin et al., 1990; Read & Perko, 1998; Sonnabend, Hughes, Giuffre, & Farrell, 1997). All studies were conducted in orthopaedic settings, two contained partial verification bias (Farin et al., 1990; Sonnabend et al., 1997) and only one used a consecutive patient series and reported clearly defined surgical diagnostic criteria for impingement (Sonnabend et al., 1997). These studies concluded that ultrasound imaging may help confirm, but not exclude the diagnosis of "impingement". However the small number of studies, methodological flaws, unknown association between observed impingement and symptoms, as well as the questionable value of surgical findings as a reference standard for a dynamic clinical diagnosis of impingement mean this technique is not suitable as a reference standard procedure for subacromial impingement.

Rotator cuff tears

A large number of studies and two systematic reviews were identified investigating the reliability and diagnostic accuracy of ultrasound imaging in identifying

rotator cuff tears. The two systematic reviews included cohort studies published from 1985 to 2001 (Dinnes et al., 2003), and 1966 to 2007 (de Jesus et al., 2009).

Both reviews were high quality, with de Jesus et al. (2009) using similar search strategies and updating the previous systematic review published by Dinnes et al. (2003). While Dinnes et al. (2003) assessed the methodological quality of the studies, de Jesus et al. (2009) used study inclusion criteria that included surgery as the only reference standard procedure, and in which imaging results were read and interpreted only by radiologists, which is more reflective of primary care practice than reports interpreted by rheumatologists or orthopaedic specialists.

De Jesus et al. (2009) also performed a meta-analysis, pooling data from studies in which contingency cell counts were extractable to provide summary sensitivity and specificity for full and partial thickness rotator cuff tears. The authors acknowledged several limitations, most notably the heterogeneous diagnostic criteria used to define partial-thickness rotator cuff tears. This review included studies conducted between 1966 and 2007, during which time significant advances have been made in ultrasound equipment technology, skill and expertise which may have contributed to the inconsistent diagnostic criteria used in studies over this period of time. The methodological quality of studies included in this meta-analysis was not formally assessed, and selection bias, clinical review bias and partial verification bias were present in a number of studies included in the pooled analysis, largely due to retrospective design. Many of these studies were included in both systematic reviews hence these limitations are likely to affect results of both systematic reviews. Despite these limitations, and technological advances that have occurred since the review was published in 2003, the results of de Jesus et al. (2009) were similar to those of Dinnes et al. (2003) for the diagnostic accuracy of ultrasound for identification of rotator cuff tears.

The diagnostic accuracy results (sensitivity and specificity) of diagnostic ultrasound imaging for rotator cuff tears from individual studies, and two meta-analyses are summarised in Table 2.7. Only individual studies in which variable high frequency (more than 7.5MHz) linear array transducers were used are included in the summary due to their superior imaging ability for rotator cuff disease (Teefey et al., 2004). For a full thickness rotator cuff tear, the majority of studies reported sensitivity and specificity of more than 0.90 (C. Y. Chang et al., 2002; Fotiadou et al., 2008; Frei, Chladek, Trc, Kopecny, & Kautzner, 2008; Milosavljevic, Elvin, & Rahme, 2005; Moosmayer, Heir,

& Smith, 2007; Zehetgruber, Lang, & Wurnig, 2002; Ziegler, 2004). The pooled sensitivity and specificity of ultrasound for a full thickness rotator cuff tear were also similarly high in the two meta-analyses despite having included older studies involving the use of low frequency transducers and non-radiology professionals (de Jesus et al., 2009; Dinnes et al., 2003). Based on the studies included in the analysis by Dinnes et al. (2003), a positive ultrasound finding of a full thickness tear increases the probability of such a tear being present from around 50% to over 90% (Dinnes et al., 2003).

Table 2.7. *Summary of Diagnostic Accuracy for Diagnostic Ultrasound Imaging*

	Individual studies	Dinnes et al. (2003)	de Jesus et al. (2009)
Any tear			
sensitivity %	66[a] to 98[b]	80	85
(95% CI)		(78 to 83)	(1.6%)
specificity %	60[c] to 98[b]	85	92
(95% CI)		(82 to 87)	(1.2%)
Full thickness tear			
sensitivity %	24[d] to 100[e]	87	92
(95% CI)		(84 to 89)	(2.1%)
specificity %	61[d] to 100[f]	96	94
(95% CI)		(94 to 97)	(1.7%)
Partial thickness tear			
sensitivity %	70[g] to 94[h]	67	67
(95% CI)		(61 to 73)	(5.9%)
specificity %	83[g] to 98[i]	94	94
(95% CI)		(92 to 96)	(1.7%)

[a](Moosmayer et al., 2007); [b](Moosmayer & Smith, 2005); [c](Ardic et al., 2006); [d](Goldberg, Bruce, Walsh, & Sonnabend, 2003); [e](Frei et al., 2008; Milosavljevic et al., 2005; Moosmayer et al., 2007); [f](C. Y. Chang et al., 2002; Cullen, Breidahl, & Janes, 2007; Fotiadou et al., 2008); [g](Iannotti et al., 2005); [h](Ziegler, 2004); [i](Milosavljevic et al., 2005)

For partial thickness tears, sensitivity was reported to be lower than for full thickness tears, but reported specificities remained high (Table 2.7) (Cullen et al., 2007; Fotiadou et al., 2008; Iannotti et al., 2005; Milosavljevic et al., 2005; Ziegler, 2004). Both de Jesus et al. (2009) and Dinnes et al. (2003) reported pooled sensitivity for partial thickness tear to be slightly lower than reports in individual studies (67%) however data from older studies that used lower frequency ultrasound equipment were included in the meta-analysis and may have reduced pooled sensitivity values. Differences between studies in categorisation of full thickness tears as either a "positive" or "negative" result for a partial thickness tear may also have affected the summary diagnostic statistics. Inconsistent ultrasound diagnostic criteria for partial thickness rotator cuff tears were identified among a group of musculoskeletal ultrasound experts in contrast to more consistent criteria for full thickness tears in the same group

(Naredo et al., 2006). Despite this, agreement between expert examiners from several countries for shoulder tendon lesions was high overall (88%) (Naredo et al., 2006).

For any type of rotator cuff tear, individual studies reported sensitivity and specificity were variable (Table 2.7) (Ardic et al., 2006; Milosavljevic et al., 2005; Moosmayer et al., 2007; Moosmayer & Smith, 2005; Teefey et al., 2004; Yen et al., 2004; Ziegler, 2004). Two studies reported both sensitivity and specificity to exceed 0.90, in which high frequency (10MHz) equipment was used and imaging was interpreted by a trained radiologist (Milosavljevic et al., 2005; Yen et al., 2004). Pooled sensitivity and specificity were moderate to high for any rotator cuff tear (de Jesus et al., 2009; Dinnes et al., 2003) (Table 2.7).

An additional search using similar search terms was undertaken to update the reviewed studies between 2007 and 2009. This identified two new studies in addition to the systematic review reference lists (Kang et al., 2009; Kelly & Fessell, 2009). One reported 3D ultrasound to have sensitivity 0.88 and specificity of 0.90 for full-thickness supraspinatus tears (Kang et al., 2009) however this study contained selection bias (all patients included were suspected of having a rotator cuff tear) and partial verification bias (only 39% received the reference standard procedure). The 3D ultrasound imaging equipment is not yet readily available, and diagnostic accuracy, as reported in this study does not appear to improve that of the pooled results reported by de Jesus et al. (2009) using conventional ultrasound equipment.

In terms of clinical decision making, identification of a partial thickness or full thickness tear has the potential to alter management decisions and evidence suggests that diagnostic ultrasound has the ability to do this when performed by an experienced operator using modern equipment.

Rotator cuff tendinopathy.

Only two studies were identified in which the diagnostic accuracy of ultrasound was investigated specifically for rotator cuff tendinopathy (Martin-Hervas et al., 2001; Naredo et al., 1999). One of the studies was conducted in an orthopaedic care setting, recruited patients from a surgical waiting list, using combined reference standards (arthroscopy and MRI) and reported the sensitivity of ultrasound to be 0.67 and specificity 0.88 for supraspinatus tendinopathy. However details relating to eligibility criteria, standardisation of reference standard procedures, diagnostic criteria (arthroscopy or MRI) and blinding were lacking from this study (Martin-Hervas et al., 2001). A higher quality study conducted in a rheumatology clinic using patients

reporting a first flare of shoulder pain reported higher diagnostic accuracy values (sensitivity 0.93, specificity 1.00) for ultrasound diagnosed supraspinatus tendinosis compared with MRI (Naredo et al., 1999). Differences in setting, recruitment, reference standard procedures and methodological quality mean these two studies cannot be directly compared and results cannot be generalised to the primary care setting. There is insufficient evidence to draw conclusions regarding the use of ultrasound as a reference standard for the identification of tendinopathy.

Calcific tendon lesions.

Only two studies were found investigating the diagnostic ultrasound for calcific tendon lesions (Kayser et al., 2005; Martin-Hervas et al., 2001). Both were conducted in orthopaedic departments recruiting patients awaiting surgery and using arthroscopy as the reference standard. The study by Kayser et al. (2005) demonstrated superior methodological quality including a larger sample size (239) (Kayser et al., 2005), and reported sensitivity of 1.00 and specificity of 0.98 for ultrasound diagnosis of calcific tendinitis by two orthopaedic surgeons. Martin-Hervas et al. (2001) also reported high sensitivity (1.00; 95% CI 62.9%, 100%), but slightly lower specificity (0.84; 95% CI 70.9%, 92.5%) for a radiologist diagnosis of supraspinatus calcific tendinitis compared with arthroscopy (Martin-Hervas et al., 2001). Based upon these results ultrasound appears to be of some value for ruling-out calcific tendon lesions, and may be of value for ruling-in calcific tendinitis in some settings but more studies are required to confirm these results.

Acromioclavicular joint pathology

Only one study was identified in which the diagnostic accuracy of ultrasound was evaluated for degenerative ACJ pathology compared with MRI (Naredo et al., 1999). This study reported sensitivity and specificity of 1.00 for ultrasound diagnosed degenerative ACJ changes using a 7.5MHz transducer, performed by a rheumatologist. While the quality of this study was high, more studies are required before ultrasound can be deemed an acceptable reference standard for degenerative ACJ pathology or other pathology affecting this joint.

Glenohumeral joint pathology

Only one study was identified investigating the diagnostic value of ultrasound for pre-operative evaluation of anterior instability including glenoid labrum tears (Hammar et al., 2001). The anterior labrum, the anterior ligamentous-capsular complex, and the presence of humeral head and glenoid rim fractures were evaluated using three dynamic

scanning approaches in 22 patients (20 to 40 years of age) with acute traumatic or recurrent anterior shoulder dislocation, and results compared with arthroscopic or arthrotomy findings. Ultrasound had a reported sensitivity of 0.88 to 0.95 and a specificity of 0.67 to 1.00 in the diagnosis of anterior labral tears. While these results appear promising for ultrasound in ruling-out an anterior labral tear, the reference standard procedure was not standardised for all participants, average time between ultrasound and the reference standard procedure was 3.4 months (maximum 8 months), the sample size was small (n=22) and while the dynamic ultrasound scan techniques were well described, reliability was not reported. More studies are required on larger sample sizes before ultrasound can be considered an acceptable reference standard for anterior labral lesions.

Similar to studies investigating diagnostic accuracy of clinical examination tests for specific shoulder pathology, all studies investigating the diagnostic accuracy of ultrasound findings were conducted in secondary care, rheumatology or radiology departments. Many samples included patients who had failed conservative management, had confirmed diagnoses or rheumatic conditions. The prevalence of conditions and spectrum of disease reported in these studies is likely to differ from that seen in primary care. The accuracy of ultrasound for specific shoulder pathology in primary care therefore remains to be evaluated.

Factors affecting the ability of ultrasound to detect shoulder pathology.

In all the studies reviewed, a number of factors were identified that may influence the reliability and subsequent diagnostic accuracy values reported for ultrasound diagnosis of shoulder pathology. Equipment specifications, including resolution and transducer frequency affect accurate visualisation of pathology, and agreement between examiners was shown to reduce agreement between examiners for tendon, bursae and joint measurements from 0.96 (0.87–0.99) (10MHz transducer) to 0.60 (0.04–0.99) with lower frequency equipment (Schmidt et al., 2004).

Diagnostic ultrasound is commonly referred to as being an "operator dependent" investigation (Read & Perko, 1998; Rutten, Jager, & Blickman, 2006; M. T. Van Holsbeeck et al., 1995). While low levels of agreement have been reported between experienced operators and less experienced operators for full-thickness rotator cuff tears (kappa 0.18 to 0.21) (O'Connor et al., 2005), recent evidence suggests that sonography of the rotator cuff is an accurate and reproducible diagnostic test when performed by experienced examiners using modern equipment for identifying full-thickness rotator

cuff tears (C. Y. Chang et al., 2002; Le Corroller et al., 2008; O'Connor et al., 2005), partial thickness tears (Cullen et al., 2007), tendon calcification, dynamic signs of impingement (Naredo et al., 2006; Schmidt et al., 2004), bursitis (Naredo et al., 2006) and abnormality of the long head of biceps tendon (O'Connor et al., 2005). These results suggest that while ultrasound may be a less operator-dependent investigation when performed by experienced staff using modern equipment, agreement may still be poor when there is marked disparity between the operators' experience levels, or when using low frequency equipment.

Validity of Magnetic Resonance Imaging

Since the 1990's there has been a proliferation of studies in which the reliability and accuracy of MRI compared with surgery was reported for identification of specific shoulder pathology. Results are summarised below.

Subacromial pathology

A large number of studies compared the diagnostic accuracy of conventional MRI or MRA with surgery for identifying tears of the rotator cuff. The accuracy of MRI and MRA in identifying rotator cuff tears was also the subject of a systematic review (Dinnes et al., 2003) and meta-analysis conducted by de Jesus et al. (2009).

Rotator cuff tears.

Conventional MRI studies report a wide variation in diagnostic accuracy for the detection of rotator cuff tears compared with surgery (Table 2.8). A number of studies reported sensitivity and specificity exceeding 0.90 for full thickness tears (Balich, Sheley, Brown, Sauser, & Quinn, 1997; Robertson et al., 1995; Traughber & Goodwin, 1992; Waldt et al., 2007). This is similar to results of the meta-analysis in which only studies that had been reported by trained radiologists were included, with reported pooled sensitivity of 0.92 (CI 2.1%) and pooled specificity 0.93 (CI 1.5%) for full thickness tears (de Jesus et al., 2009). Magnetic resonance imaging sensitivity of partial thickness tears demonstrated more variation in individual studies, however pooled specificity for partial thickness tears remained high (Table 2.8). For identification of either a full thickness or partial thickness tear, pooled sensitivity and specificity of MRI were 0.87 (CI 2.6%) and 0.82 (CI 3.5%) respectively (de Jesus et al., 2009). Sensitivity and specificity of MRI for any rotator cuff tear were moderate to high in individual studies, as well as the meta-analysis (Table 2.8) (de Jesus et al., 2009).

Consistently lower sensitivities were reported for community radiologists for rotator cuff tears (compared with trained musculoskeletal radiologists) (Wnorowski et al., 1997). Dinnes et al. (2003) included studies that used a range of reference standards that were interpreted by a variety of health professionals including orthopaedic specialists, radiologists and rheumatologists. They were unable to identify clear sources of heterogeneity between studies despite subgrouping them according to MRI sequences and imaging planes, publication year, age of participants, prevalence of MRI findings, study design, reference standard used and presence of bias. However they did not subgroup results according to observer training or experience. Despite this their pooled sensitivity and specificity results were similar to those of de Jesus et al. (2009) in which only studies where images had been reported by trained radiologists were included.

Subacromial bursa and tendon pathology.

Few studies were identified investigating the diagnostic accuracy of MRI or MRA for bursal pathology (Ardic et al., 2006; Farley, Neumann, Steinbach, Jahnke, & Petersen, 1992) or tendinopathy (Farley et al., 1992; Iannotti et al., 1991; Robertson et al., 1995). Although the accuracy of MRI for surgically identified bursal pathology was not specifically reported, the MRI appearance of bursal effusion was reported to be 93% sensitive for a full thickness rotator cuff tear (Farley et al., 1992). Conventional MRI was also found to be superior to diagnostic ultrasound for identifying subacromial bursal effusion or hypertrophy, (p<0.01) however MRI was used as the reference standard in this study (Ardic et al., 2006).

Variable sensitivity (0.13 to 0.82) and moderate specificity (0.73 to 0.85) have been reported for MRI detection of 'tendinitis' (Iannotti et al., 1991; Robertson et al., 1995). In an older study, poor inter-observer agreement was reported among radiologists for the MRI diagnosis of tendinitis (kappa 0.14 to 0.27) (Robertson et al., 1995). Abnormal signal (tendinitis) was reported in 27/31 (87%) of participants with confirmed full thickness rotator cuff tear, however diagnostic accuracy statistics were not reported (Farley et al., 1992).

Table 2.8. *Summary of Diagnostic Accuracy of MRI and MRA for Rotator Cuff Tears*

Type of tear	Conventional MRI			MRA		
	Individual studies	Dinnes et al. (2003)	de Jesus et al. (2009)	Individual studies	Dinnes et al. (2003)	de Jesus et al. (2009)
Sensitivity % (95% CI)						
any tear	71[a] to 92[b]	83 (79, 86)	86 (1.8%)	71[c] to 100[d]	88 (80, 93)	92 (1.8%)
full thickness tears	56[e] to 100[f]	89 (0.86, 0.92)	92 (2.1%)	91 to 100[g]	95 (82, 98)	95 (2.7%)
partial thickness tears	0[e] to 97[h]	44 (36, 51)	64 (6.2%)	80[i]	62 (40, 80)	86 (4.9%)
Specificity % (95% CI)						
any tear	52[e] to 100[j]	86 (83, 88)	90 (1.2%)	78 to 89[d]	83 (78, 89)	96 (0.7%)
full thickness tears	73[e] to 100[k]	93 (91, 95)	93 (1.5%)	79 to 100[g]	93 (84, 97)	99 (0.7%)
partial thickness tears	68[e] to 93[l]	90 (87, 92)	92 (1.8%)	97[i]	92 (83, 97)	99 (0.7%)

[a](Traughber & Goodwin, 1992); [b](Burk Jr et al., 1989; Frei et al., 2008); [c](Waldt et al., 2007); [d](Yağcı et al., 2001); [e](Wnorowski, Levinsohn, Chamberlain, & McAndrew, 1997); [f](Fotiadou et al., 2008; Robertson et al., 1995; Traughber & Goodwin, 1992); [g](Binkert, Zanetti, Gerber, & Hodler, 2001); [h](Fotiadou et al., 2008); [i](Waldt et al., 2007); [j](Frei et al., 2008) [k](Fotiadou et al., 2008; Frei et al., 2008; Traughber & Goodwin, 1992); [l](Robertson et al., 1995).

Validity of Magnetic Resonance Arthrography and Comparison with MRI

Since the development of conventional MRI technology, there have also been relatively rapid advances in technical procedures involving contrast enhancement of specific shoulder structures. Magnetic resonance arthrography (MRA) was developed, in which contrast is administered intravenously (indirect MRA) or directly into the glenohumeral joint (direct MRA) prior to obtaining magnetic resonance images. With developments in these techniques, many more studies have been published in which the accuracy of MRA has been directly compared with surgical findings, and the accuracy of both direct and indirect MRA procedures has also been compared. The majority of studies investigated and compared the accuracy of both MRI and MRA for rotator cuff tears or glenoid labrum pathology.

One general review (Shahabpour et al., 2008) and two systematic reviews with meta-analysis (de Jesus et al., 2009; Dinnes et al., 2003) were also identified comparing the accuracy of MRI and MRA with arthroscopically or surgically diagnosed shoulder pathology. The main findings from individual studies, the systematic reviews and meta-analysis for MRI and MRA are summarised in the following sections for specific shoulder pathologies.

Subacromial pathology

A small number of studies were identified in which MRA was compared with surgery for identification of rotator cuff tears (Binkert et al., 2001; Funke, Kopka, Vosshenrich, Oestmann, & Grabbe, 1996; Hodler et al., 1992; Loew, Kreitner, Runkel, Zoellner, & Thelen, 2000; Pfirrmann, Zanetti, Weishaupt, Gerber, & Hodler, 1999; Wagner, Schweitzer, Morrison, Fenlin J.M, & Bartolozzi, 2002; Waldt et al., 2007; Yağcı et al., 2001) (Table 2.8). Studies generally reported variable sensitivity for partial and full thickness tears (0.00 to 1.00), with more consistent levels of specificity for these pathologies (0.68 to 1.00). Only one study investigating MRA for partial thickness tears was identified and this study found that the majority of false-negative and false-positive results (78%) occurred for small articular surface tears (Waldt et al., 2007).

The two systematic reviews and meta-analysis did not differentiate between results of studies in which direct and indirect MRA techniques were used (de Jesus et al., 2009; Dinnes et al., 2003). Although few studies had directly compared direct and indirect MRA for detection of shoulder pathology, there was some evidence that direct MRA is more accurate (100%) than indirect MRA (83%) in identifying rotator cuff tears (Wagner et al., 2002).

When conventional MRI was compared with MRA for the ability to detect rotator cuff tears, reported sensitivity in individual studies, as well as pooled sensitivity of MRA for any rotator cuff tear (including full thickness and partial thickness tears) was higher than reported sensitivities for conventional MRI (Table 2.8). Based on summary likelihood ratios and average prevalence of rotator cuff tears, Dinnes et al. (2003) calculated that a negative MRI reduced post-test probabilities of a full thickness rotator cuff tear from 30% to under 10%, whereas a negative MRA reduced the probability from 36% to just over 5% (Dinnes et al., 2003). A positive MRA finding of full-thickness tear increased the probability of such a tear being present from 36% to over 80% (Dinnes et al., 2003). These results indicate a trend of improved sensitivity of MRA compared with MRI for a tear of the rotator cuff with similar specificity for both procedures.

Acromioclavicular joint pathology

The MRA procedure is typically used for detecting GHJ pathology, and no studies were identified in which arthrographic procedures involving injection of contrast into the ACJ were used to detect pathology affecting this joint. No studies reported the accuracy of either MRI or MRA for detecting ACJ pathology compared with surgical observations.

Glenohumeral joint pathology

Several studies were identified investigating the diagnostic accuracy of MRI and MRA for capsuloligamentous and glenoid labrum lesions (Bencardino et al., 2000; Borrero, Casagranda, Towers, & Bradley, 2010; Chandnani et al., 1995; Connell, Potter, Wickiewicz, Altchek, & Warren, 1999; Cvitanic et al., 1997; De Maeseneer et al., 2000; Dinauer, Flemming, Murphy, & Doukas, 2007; Gusmer et al., 1996; Iannotti et al., 1991; Jee et al., 2001; W Jin, K.N Ryu, S.H Kwon, Y.G Rhee, & D.M Yang, 2006; Legan et al., 1991; Monu, Pope Jr, Chabon, & Vanarthos, 1994; W. E. Palmer & Caslowitz, 1995; Şahin & Demirtaş, 2006; Schreinemachers, Van Der Hulst, Willems, Bipat, & Van Der Woude, 2009; S. J. Snyder, Banas, & Karzel, 1995; Steinbach, Palmer, & Schweitzer, 2002; Tuite et al., 2005; Wagner et al., 2002; Waldt et al., 2005; Waldt et al., 2004).

The role of MRA for imaging the glenoid labrum appears to be contentious, however the majority of literature reports indicate superior diagnostic ability of MRA compared with MRI for labral tears, particularly in athletic populations (Connell et al., 1999; Gusmer et al., 1996; W. Jin, K. N. Ryu, S. H. Kwon, Y. G. Rhee, & D. M. Yang,

2006; Legan et al., 1991; Monu et al., 1994; S. J. Snyder et al., 1995; Steinbach et al., 2002). Recent studies support these claims, reporting sensitivities between 0.82 to 1.00 for MRA compared with 0.66 to 0.85 for MRI (Applegate et al., 2004; Bencardino et al., 2000; Dinauer et al., 2007; Jee et al., 2001; Wagner et al., 2002; Waldt et al., 2004). However less difference in specificity was reported between the two procedures (MRA 0.71 to 0.98; MRI0.75 to 0.83).

The weight of evidence also favours the use of the ABER position during MRA for more accurate identification of capsulolabral and ligamentous injury compared with the standard 'arm by the side' position (Borrero et al., 2010; Cvitanic et al., 1997). Compared with conventional MRA (sensitivity 0.48; specificity 0.91) MRA obtained with the patient in the ABER position increased both sensitivity and specificity to 0.89 and 0.95 respectively (Cvitanic et al., 1997). When reviewed together with the conventional MRA images, sensitivity and specificity further improved to 0.96 and 0.97 respectively.

Prospective studies have reported high sensitivities (0.89 to 0.92) and specificities (0.91 and 0.92) of direct MRA for detection of SLAP lesions and labral abnormalities (Bencardino et al., 2000; W. E. Palmer & Caslowitz, 1995), and two retrospective studies also reported similar results for labro-ligamentous lesions (sensitivity 0.82 to 0.92; specificity 0.69 to 0.92) (Jee et al., 2001; Waldt et al., 2005). High sensitivity and specificity (0.80 to 1.00) have also been reported for detection of lesions of the glenohumeral ligaments (Chandnani et al., 1995).

Overall, the majority of diagnostic accuracy studies reviewed contained various sources of bias common to other diagnostic studies including selection bias similar to those reported in diagnostic accuracy studies involving physical examination tests. These included non-consecutive patient series, retrospective design, lack of reported diagnostic criteria for the reference standard, inconsistently applied reference standard, poorly described reference standard, partial verification bias or long durations between the index test and the reference standard procedure (surgery). These methodological limitations mean results of these studies should be interpreted with caution, particularly when applying results to different populations such as primary health care.

Key Findings

- Many studies contained sources of bias and variation.

- Equipment specification varied considerably between studies, and technological advances over time mean results of earlier studies may no longer reflect the true ability of diagnostic imaging procedures to accurately detect shoulder pathology.

Diagnostic Ultrasound

- Ultrasound diagnosis of subacromial 'bursitis' demonstrates variable agreement between examiners, lack of consistent diagnostic criteria and different interpretations of findings as pathological or normal.

- Preliminary reports support accuracy of ultrasound for ACJ pathology, tendinopathy and calcific lesions. There is insufficient research to draw firm conclusions regarding the diagnostic accuracy of ultrasound for the diagnosis of glenoid labrum lesions.

- Ultrasound demonstrates acceptable reliability and diagnostic accuracy compared with surgical findings for identifying full-thickness tears, and for ruling-in partial-thickness when performed by experienced operators using modern equipment.

Magnetic Resonance Imaging and Magnetic Resonance Arthrography

- Direct MRA accurately identifies full thickness rotator cuff tears and is highly sensitive for partial thickness rotator cuff tears.

- Direct MRA accurately identifies glenoid labral tears and capsuloligamentous lesions when the ABER position is used.

- There is insufficient evidence to draw conclusions regarding the diagnostic accuracy of MRI or MRA for subacromial bursal pathology or tendinopathy.

CHAPTER THREE

RELIABILITY OF THE CLINICAL EXAMINATION

Preface

This chapter relates to Specific Aim 1 of the thesis:

To evaluate the reliability of clinical examination tests being considered for inclusion in a standardised clinical examination.

In relation to this aim, a reliability study was undertaken to evaluate the reliability of clinical examination tests being considered for use in the diagnostic accuracy study. The reliability of physical examination tests, including measures of range of motion (ROM), strength (peak isometric force), and symptom responses associated with ROM, resisted tests and orthopaedic special tests (OST) was investigated. Results are presented in a series of four manuscripts.

The first manuscript reports reliability results for measures of range of motion (ROM) and strength (peak isometric muscle force). This manuscript was published in Manual Therapy (2011) (Cadogan, Laslett, et al., 2011a). The second manuscript reports reliability results for results of orthopaedic special tests (OST) used in the assessment of shoulder pain, and this manuscript was also published in Manual Therapy (2011) (Cadogan, Laslett, et al., 2011b). The third manuscript presents unpublished results for the reliability of symptom responses associated with ROM and resisted muscle tests.

Methodological development was then undertaken to refine test procedures in an attempt to reduce measurement variability associated with measures of ROM and peak isometric muscle force, prior to inclusion of these tests in the diagnostic accuracy study. The scope of the PhD did not allow for additional extensive reliability testing following these methodological developments, however intraexaminer reliability was re-evaluated in a small sample of volunteers. These (unpublished) results are presented in the final section in this chapter.

3.1. Reliability of a New Hand-Held Dynamometer in Measuring Shoulder Range of Motion and Strength

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Williams, M. (2011). Reliability of a new hand-held dynamometer in measuring shoulder range of motion and strength. *Manual Therapy, 16*(1), 97-101.

A copy of the published manuscript is included in Appendix 4 (p270).

Abstract

Acceptable reliability is a prerequisite for inclusion of physical examination tests in clinical examinations of the painful shoulder. The aim of this study was to establish the intraexaminer and interexaminer reliability of measures of shoulder range of motion (ROM) and muscle force using a new hand-held dynamometer with the ability to standardise overpressure force during passive ROM tests. Forty consecutive participants with shoulder pain were recruited, and tests were performed by two physiotherapists. Tests included active ROM elevation, passive ROM glenohumeral abduction and external rotation and resisted abduction and external rotation. All tests demonstrated high levels of intraexaminer reliability (ICC 0.85-0.99; LOA 6-24 degrees and 1.1-7.0kg). Highest levels of interexaminer reliability were observed for measures of active ROM flexion (ICC 0.88-0.95; LOA 14-22 degrees). Passive ROM tests demonstrated 'moderate' to 'substantial' interexaminer reliability (ICC 0.45-0.62; LOA 25-34 degrees). The ICCs for resisted tests ranged from 0.68-0.84, and LOA ranged from 3.2-8.5kg. Active ROM flexion demonstrated high levels of both intra- and interexaminer reliability. Measures of passive ROM and peak isometric force demonstrated acceptable levels of intraexaminer reliability.

Introduction

Shoulder pain is a common complaint resulting in significant pain, functional disability and loss of quality of life (Lin et al., 2005; J. C. MacDermid, Ramos, Drosdowech, Faber, & Patterson, 2004; Turner-Bowker, Bayliss, & Ware, 2003). The diagnosis of shoulder pain involves a clinical examination which typically consists of a variety of physical examination tests and associated measures including active and passive range of motion (ROM), and resisted muscle tests. The results of these tests

including measures of active shoulder elevation and passive ROM abduction and external rotation are commonly used for diagnostic classification, and in the assessment of functional impairment (Constant & Murley, 1987; Davis, 1998; Harrington et al., 1998; J. C. MacDermid et al., 2007). Reliable measurements are required if these classifications are to be consistently applied.

Although few studies have directly compared reliability between active and passive ROM of the shoulder, more variability has been reported in measures of passive ROM (ICC 0.26-0.90) versus active ROM (ICC 0.49-0.88) (K. Hayes, Walton, Szomor, & Murrell, 2001; Hoving et al., 2002; Riddle, Rothstein, & Lamb, 1987; Terwee et al., 2005). A common explanation for this variability in measures of passive ROM is the inability of the examiner to standardise the amount of overpressure applied at the end range of motion (Boone et al., 1978; Gajdosik & Bohannon, 1987; K. Hayes et al., 2001; Lea & Gerhardt, 1995). A new hand-held dynamometer (HHD) (Industrial Research Ltd) has been developed that has the ability to simultaneously measure both angle and force. This feature enables the standardisation of overpressure force at end range of motion. Whether this feature would reduce measurement variability and improve reliability during measures of passive ROM has not been tested to date.

Measures of muscle strength are used in the diagnostic process to assess muscle integrity and to determine the level of any strength deficits (Constant & Murley, 1987; Cyriax, 1982). Hand-held dynamometry has demonstrated higher sensitivity, and interexaminer reliability than manual muscle testing in identifying strength deficits of the rotator cuff (Ellenbecker, 1996; K. Hayes, Walton, Szomor, & Murrell, 2002; Leggin, Neuman, Iannotti, Williams, & Thompson, 1996; Tyler et al., 2005). Hand-held dynamometry therefore provides an advantage over manual muscle testing for the accurate clinical assessment of isometric muscle strength. The reliability of the new HHD needs to be established for measurements of muscle strength before it can be used for this purpose in the clinical setting.

Thus, the aim of this study was to evaluate the intra- and interexaminer reliability of a new HHD in measuring ROM and isometric muscle strength of the symptomatic shoulder. Whether reliability of measures of passive ROM could be improved by standardizing the amount of overpressure force applied at end range of motion was specifically investigated.

Methods

Participants

Forty consecutive participants with shoulder pain were recruited from local physiotherapy practices. Participants were included if they were over 18 years of age and were currently experiencing shoulder pain. Participants were excluded if they had pain referred from a source other than the shoulder, fractures or dislocations around the shoulder joint, or were suffering known systemic inflammatory disease. Ethical approval was gained from the Ministry of Health Ethics Committee.

Hand-Held Dynamometer

The Industrial Research Ltd hand-held dynamometer (HHD) has the ability to simultaneously measure angle (degrees) and force (kg) (Figure 3.1A). This feature enables standardisation of load applied at end of range of motion during passive movement testing by selecting a force level at which an audible alarm is produced. Force is recorded from an in-built force transducer attached to the HHD, and peak force (kg) is displayed. The HHD is gravity dependent, and indicates range of motion on a 360 degree scale with reference to the vertical plane. The HHD records absolute range of motion calibrated from a zero position, and final range of motion. The relative range is displayed on the unit being calculated by subtracting the initial starting range from the final range of motion. Each examiner used a separate HHD, and both were calibrated on the first day of data collection to ±1 degree and to within ±0.1kg of force.
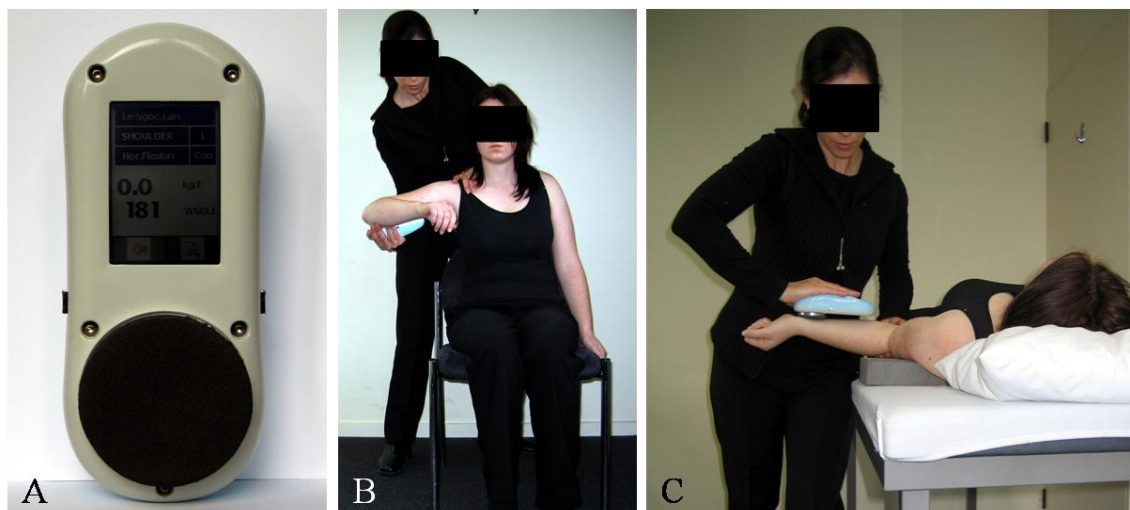
Figure 3.1. Hand-held dynamometer used for measuring range of motion and peak isometric muscle force. Figure shows: A) hand-held dynamometer; B) test procedures for glenohumeral abduction and C) test procedure for passive external rotation (at 0° abduction).

Procedures

Prior to the study, the examiners underwent 4 sessions of familiarization training with the HHD. Seven physical examination tests were performed. Range of motion was measured during active ROM elevation (through flexion), passive ROM glenohumeral abduction and external rotation (performed at 0 degrees of abduction). Peak isometric force (kg) was measured during resisted abduction and external rotation on both the affected and unaffected sides.

The physical examination tests were performed on the same day by two experienced physiotherapists (19 and 38 years experience). The examiners were blinded to each other's results. Examiner sequence and the order of tests were randomized using a random sequence generator for each participant and each examiner.

Range of motion tests.

For active ROM elevation through flexion, each participant stood with their back against the wall to prevent compensatory movement of the trunk. The HHD was aligned along the long axis of the humerus and three trials were performed. Each trial was followed by approximately 30 seconds of rest. For tests of passive ROM glenohumeral abduction, participants were positioned as in Figure 3.1B. The examiner applied firm downward pressure over the acromion while the participants' arm was guided into abduction in the scapula plane. To compensate for variation in participant limb mass when raising the arm against gravity, a standardised force equivalent to 6% of the participants' body mass was programmed into the HHD and when this force level was reached, the sound of the audible alarm was used as the criteria for end range. The 6% body mass level was determined following pilot testing on a sample of patients with shoulder pain. This force level was consistently found to be required in order to reach end range of motion without excessive discomfort to the patient. Measurement of shoulder external rotation is shown in Figure 3.1C. A load was required that would overcome the mass of the arm to allow end range of motion to be achieved, without causing excessive discomfort to the participant. Average upper limb mass is approximately 3-4% of total body mass (Clarys & Marfell-Jones, 1986), and based upon results of pilot testing a standardised load of 3kg was selected as it appeared to consistently fulfil end-range and comfort criteria for all participants.

Resisted muscle tests.

Resisted abduction was performed with the participant sitting with the arm in approximately 10 degrees of abduction in the plane of the scapula. The HHD was

placed immediately proximal to the lateral epicondyle of the humerus. Resisted external rotation was performed with the participant sitting with the forearm in neutral, the arm by the side, exerting a slight adduction force against the examiner's hand while simultaneously exerting an external rotation force against the HHD held against the distal forearm. Three maximal isometric contractions were performed, the duration of which was approximately 6-7 seconds and 30 seconds rest was given between trials. Participants were instructed to hold the contraction against maximal examiner pressure and peak isometric muscle force was recorded.

Statistical Methods

Relative reliability was assessed using single-measure intraclass correlation coefficients ($ICC_{2,1}$) and associated 95% confidence intervals (CI) (two-way random effects model -absolute agreement). For intraexaminer reliability, data from the first trial was compared with data from the second trial, and the mean of three trials for both examiners. For interexaminer reliability, data from a single trial (first trial (ROM) or the peak force trial (resisted tests)), and the single trial compared with the mean of three trials were used for the analysis. A one-way ANOVA was used to ascertain any differences between trials (intraexaminer reliability) and between examiners (interexaminer reliability) with the level of significance set at $p=0.05$.

Absolute reliability was determined by calculating the mean difference between measures and the associated 95% CI for the mean difference, as well as limits of agreement (LOA) according to the Bland and Altman method of assessing agreement (mean difference between examiners $\pm$ 1.96 SD_{diff}) (Bland & Altman, 1986).

Reliability values were interpreted according to the guidelines of Landis and Koch (1977); 0.00-0.20 slight; 0.21-0.40 fair; 0.41-0.60 moderate; 0.61-0.80 substantial; 0.81-1.00 almost perfect (Landis & Koch, 1977).

Results

Demographic characteristics of the participants are provided in Table 3.1. Twenty one participants were examined first by examiner 1 and 19 by examiner 2.

Table 3.1. *Summary of Participant Characteristics*

Characteristics		Number	%
Gender	male	23	58
	female	17	42
Affected Side	dominant	24	60
	non-dominant	11	28
	bilateral	5	13
		Mean	Range
Age (years)		49	18-77
Height (cm)		171	157-189
Weight (kg)		80	53-102
Duration of symptoms (months)		48	<1-325
Pain severity in previous 24 hours (11 point VAS)		3.6	0-7

Intraexaminer Reliability

Mean differences between Trial 1 and 2, and between Trial 1 and the mean of three trials for both examiners, LOA and ICC values are presented in Table 3.2. For measures of ROM, ICCs ranged from 0.85 (passive ROM abduction) to 0.99 (active ROM elevation). Limits of agreement ranged from ±6 degrees (active ROM elevation) to ±24 degrees (passive ROM abduction). For resisted tests ICCs ranged from 0.91-0.99, and LOA ranged from ±1.1kg (resisted external rotation -unaffected side) to ±7.0kg (resisted abduction – affected side). The results of comparisons between Trial 1 and the mean of three trials showed consistently higher levels of reliability and agreement than comparisons between single trials (trial 1 and 2) for all tests. A significant difference between the mean of the trials ($p<0.05$) was identified for resisted external rotation for examiner 2 (0.1-0.2kg) (Table 3.2).

Interexaminer Reliability

Mean differences between examiners, LOA and ICC values for interexaminer reliability are presented in Table 3.3. For measures of ROM, ICCs ranged from 0.45 (passive ROM abduction) to 0.95 (active ROM elevation – mean of three trials). Widest 95% CI were observed for measures of passive ROM. Limits of agreement ranged from 14 degrees (active ROM elevation – mean of 3 trials) to 34 degrees (passive ROM abduction – 1st trial). For resisted tests, ICC values ranged from 0.68-0.84, with the widest 95% CI recorded for resisted external rotation (unaffected side). Limits of agreement ranged from 3.2kg (resisted external rotation – mean of 3 trials) to 8.5kg (resisted abduction –unaffected side).

Table 3.2. *Intraexaminer Reliability for Measures of Range of Motion and Peak Isometric Force*

Physical Examination test		Examiner 1 Mean (Range)	Examiner 1 Mean difference between trials $\pm$ LOA (95% CI)[a]	Examiner 1 ICC (95% CI)	Examiner 2 Mean (Range)	Examiner 2 Mean difference between trials $\pm$ LOA (95% CI)[a]	Examiner 2 ICC (95% CI)
Range of motion (degrees)							
Active elevation	Single[b]	150 (84-181)	0.5 $\pm$ 19 (-2.5, 3.5)	0.92 (0.88, 0.96)	152 (94-182)	0.8 $\pm$ 9 (-2.1, 0.6)	0.98 (0.97, 0.99)
	Mean of 3 trials[c]	150 (88-180)	0.7 $\pm$ 14 (-1.6, 3.0)	0.96 (0.92, 0.98)	153 (95-180)	0.2 $\pm$ 6 (-1.0, 0.7)	0.99 (0.99, 1.0)
Passive abduction	Single	76 (35-126)	0.4 $\pm$ 24 (-4.2, 3.5)	0.85 (0.76, 0.91)	86 (44-128)	0.8 $\pm$ 14 (-1.4, 2.9)	0.91 (0.85, 0.95)
	Mean of 3 trials	77 (36-116)	0.5 $\pm$ 13 (-1.6, 2.6)	0.94 (0.88, 0.97)	86 (45-119)	0.5 $\pm$ 9 (-1.0, 1.9)	0.96 (0.93, 0.98
Passive external rotation	Single	46 (8-126)	1.0 $\pm$ 12 (-2.9, 0.9)	0.89 (0.83, 0.94)	61 (9-117)	2.0 $\pm$ 13 (-4.1, 0.2)	0.95 (0.92, 0.97)
	Mean of 3 trials	45 (8-106)	1.0 $\pm$ 10 (-2.6, 0.6)	0.96 (0.93, 0.98)	61 (10-112)	0.8 $\pm$ 8 (-2.1, 0.5)	0.98 (0.97, 0.99)
Peak isometric force (kg)							
Resisted abduction (AS)[c]	Single	20.0 (4.9-37.2)	0.8 $\pm$ 7.0 (-2.0, 0.3)	0.91 (0.85, 0.95)	17.0 (5.5-29.9)	0.1 $\pm$ 3.6 (-0.5, 0.7)	0.95 (0.92, 0.98)
	Mean of 3 trials	19.6 (5.0-33.1)	0.4 $\pm$ 4.1 (-1.0, 0.3)	0.96 (0.93, 0.98)	17.3 (6.4-30.7)	0.3 $\pm$ 2.2 (-0.1, 0.6)	0.98 (0.97, 0.99)
Resisted abduction (US)[d]	Single	21.0 (8.5-34.0)	0.4 $\pm$ 4.5 (-1.2, 0.3)	0.95 (0.91, 0.97)	19.0 (7.7-34.3)	0.2 $\pm$ 3.8 (-0.4, 0.8)	0.95 (0.92, 0.97)
	Mean of 3 trials	21.3 (9.0-34.2)	0.2 $\pm$ 3.0 (-0.7, 0.4)	0.98 (0.96, 0.99)	19.3 (7.9-33.5)	0.2 $\pm$ 2.7 (-0.2, 0.7)	0.98 (0.96, 0.99)
Resisted external rotation (AS)	Single	11.0 (4.6-19.7)	0.1 $\pm$ 1.7 (-0.2, 0.3)	0.96 (0.94, 0.98)	13.0 (4.7-20.4)	0.6 $\pm$ 3.2 (0.1, 1.1)	0.91 (0.85, 0.95)*
	Mean of 3 trials	11.3 (5.1-19.7)	0.0 $\pm$ 1.1 (-0.2, 0.1)	0.99 (0.97, 0.99)	12.8 (5.2-20.4)	0.5 $\pm$ 2.0 (0.1, 0.8)	0.96 (0.90, 0.98)*
Resisted external rotation (US)	Single	12.0 (5.3-18.6)	0.3 $\pm$ 1.9 (-0.6, 0.0)	0.94 (0.90, 0.97)	14.0 (5.6-22.3)	0.7 $\pm$ 3.0 (0.2, 1.2)	0.93 (0.87, 0.96)*
	Mean of 3 trials	11.9 (5.9-17.9)	0.2 $\pm$ 1.4 (-0.4, 0.0)	0.98 (0.95, 0.99)	13.9 (5.8-20.4)	0.4 $\pm$ 1.7 (0.1, 0.7)	0.97 (0.92, 0.98)*

Abbreviations: LOA, limits of agreement; CI, confidence interval; AS, affected side; US, unaffected side
[a] 95% CI for mean difference between trials
[b] Trial 1 compared to Trial 2
[c] Trial 1 compared to mean of 3 trials
* One-way ANOVA significant difference between trials ($p<0.05$)

Table 3.3. *Interexaminer Reliability for Measures of Range of Motion and Peak Isometric Force*

Test		Mean (range)	Mean diff $\pm$ LOA (95% CI)[a]	ICC (95% CI)
Range of motion (degrees)				
Active elevation	1st trial	151 (84 - 182)	3.0 $\pm$ 22 (-0.5, 6.6)	0.88 (0.78, 0.93)
	mean of 3 trials	152 (88 - 180)	2.1 $\pm$ 14 (-0.1, 4.4)	0.95 (0.90, 0.97)
Passive abduction	1st trial	81 (39 - 126)	9.0 $\pm$ 34 (3.4, 14.6)	0.45 (0.15, 0.67)*
	mean of 3 trials	73 (36 - 119)	8.9 $\pm$ 30 (4.0, 13.9)	0.49 (0.17, 0.71)*
Passive external rotation	1st trial	54 (9 - 126)	15.3 $\pm$ 31 (10.2, 20.4)	0.58 (0.04, 0.81)*
	mean of 3 trials	53 (8 - 112)	15.5 $\pm$ 30 (11.4, 19.5)	0.62 (-0.04, 0.85)*
Peak isometric force (kg)				
Abduction (AS)	peak force trial	19.8 (5.4 - 37.2)	-2.9 $\pm$ 7.3 (-4.2, -1.8)	0.81 (0.42, 0.92)*
	mean of 3 trials	18.5 (5.0 - 33.1)	-2.4 $\pm$ 6.3 (-3.4, -1.3)	0.84 (0.54, 0.93)*
Abduction (US)	peak force trial	21.5 (8.4 - 36.7)	-2.5 $\pm$ 8.5 (-4.0, -1.0)	0.77 (0.49, 0.89)*
	mean of 3 trials	20.3 (8.0 - 34.1)	-2.5 $\pm$ 7.6 (-3.8, -1.2)	0.77 (0.46, 0.89)*
External rotation (AS)	peak force trial	12.7 (5.5 - 21.1)	1.8 $\pm$ 4.4 (1.1, 2.6)	0.69 (0.23, 0.87)*
	mean of 3 trials	12.1 (5.1 - 20.4)	1.5 $\pm$ 4.0 (0.9, 2.2)	0.74 (0.36, 0.89)*
External rotation (US)	peak force trial	13.6 (6.1 - 22.3)	2.4 $\pm$ 3.6 (1.8, 3.1)	0.68 (-0.05, 0.89)*
	mean of 3 trials	12.9 (5.8 - 20.4)	2.2 $\pm$ 3.2 (1.7, 2.8)	0.70 (-0.04, 0.90)*

Abbreviations: LOA, limits of agreement; CI, confidence interval; AS, affected side; US, unaffected side
[a]95% confidence interval for mean difference between examiners
One-way ANOVA significant difference between examiners ($p<0.05$)

Zero was not contained within the 95% CI for the mean difference between examiners for any of the passive ROM or resisted muscle tests suggesting the presence of systematic bias. One-way ANOVA results also indicated significant differences between examiners for all passive ROM (1-8 degrees) and resisted muscle tests (0.6-1.3kg) (Table 3.3).

Discussion

Intraexaminer Reliability

All measures of ROM and peak isometric force used in this study demonstrated clinically acceptable levels of intraexaminer reliability (ICC 0.85-0.99). The results were higher than previous intraexaminer reliability results for measures of active ROM elevation through flexion (ICC 0.49-0.88) (K. Hayes et al., 2001; Hoving et al., 2002; Terwee et al., 2005), passive ROM abduction (ICC 0.58-0.67) and passive ROM external rotation (ICC 0.60-0.73) (K. Hayes et al., 2001; Terwee et al., 2005). However previous studies compared measurements on two separate occasions. In the present study, consecutive trials were conducted on one occasion to assess the number of trials required for satisfactory reliability (ICC >0.80). Intraexaminer reliability results for peak isometric force measures were similar to those of Hayes et al. (2002) and Leggin et al. (1996) for resisted abduction (ICC 0.84-0.96) and external rotation (0.89-0.95). Differences between trials for examiner 2 during resisted external rotation may be due to increasing familiarity of the participant with the test with subsequent trials, or alterations in examiner technique following performance of the initial trial. In summary, for the purposes of assessing diagnostic criteria and physical impairments, the tests used in this study demonstrated acceptable intraexaminer reliability within a single clinical session.

Interexaminer Reliability

Range of motion.

Measures of active ROM elevation through flexion using the new HHD reached ICC values in excess of 0.80, and 95% limits of agreement between 14-22 degrees. These results demonstrate improved reliability compared with previous results for measures of active ROM elevation through flexion (ICC 0.65) (Hoving et al., 2002), and LOA (27-36 degrees) (Triffitt, Wildin, & Hajioff, 1999).

Despite the ability to standardise overpressure force in the current study, interexaminer reliability for measures of passive ROM of the shoulder were lower than previously reported (ICC 0.64 -0.92, SEM 7.5 – 14 degrees) (Boone et al., 1978; K. Hayes et al., 2001; J.C MacDermid, Chesworth, Patterson, & Roth, 1999; Riddle et al., 1987). There are several possible explanations for this finding. This is a new device, and the amount of familiarization time required may have been underestimated. More experience with the HHD may reduce measurement variability resulting from subtle changes in planar angulation of the HHD. Factors relating to standardisation of the test procedures including participant positioning and instructions to participants may have also affected reliability. However care was taken to ensure that errors associated with these factors were addressed. The systematic error present in the results of passive ROM tests suggests an examiner source of error was present despite the care being taken to standardise procedures prior to the study.

A force relative to body weight (6%) was used as the criteria for end ROM during passive glenohumeral abduction to overcome the weight of the limb as it was lifted against gravity. Measures of passive external rotation did not require moving the limb against gravity, and an absolute force level of 3kg was selected as the criteria for end-range. No relationship between body weight and end ROM was identified during pilot testing, and this amount of load appeared to consistently achieve end range of motion for all participants. This is not surprising given the average mass of the arm in this study was approximately 2.4-3.2kg (3-4% of total body mass) (Clarys & Marfell-Jones, 1986) and the 3kg load was likely to move the arm to its final end range of motion.

Resisted muscle tests.

The results of the present study fall within the range of previously reported interexaminer reliability results for peak isometric muscle force during resisted abduction (ICC 0.79-0.92) (K. Hayes et al., 2002; Leggin et al., 1996) and slightly lower than previously reported for resisted external rotation (ICC 0.82-0.94) (K. Hayes et al., 2002; K. W. Hayes & Petersen, 2003; Leggin et al., 1996). Peak isometric force measures during resisted abduction (affected side) in the present study demonstrated high levels of reliability, however LOA indicate that 95% of measurements between examiners would lie within a range either 7.3kg higher or lower than the other examiner, and the lower CI for the ICC values indicated only 'moderate' reliability. This level of variability suggests measurements between examiners should be interpreted with caution.

Other peak isometric force measures also demonstrated wide LOA and while the ICC values were slightly higher for the mean of three trials compared with single trials during resisted muscle tests, lower 95% CI were poor (-0.5 to 0.36) indicating caution should also be used when interpreting these results. Confidence intervals for the mean difference between examiner measures (zero not contained in the confidence intervals) and ANOVA results also indicate systematic and significant differences between examiners for these measures (Table 3.3). This may be due to the known limitations of the use of HHD, including examiner strength (Wadsworth, Nielsen, Corcoran, Phillips, & Sannes, 1992), systematic differences in test procedures between examiners or a change in symptom severity with repeated testing. Caution is recommended when interpreting peak isometric muscle force measures during both resisted abduction and external rotation.

Conclusion

Measures of active ROM elevation (flexion) obtained using the HHD were reliable within- and between examiners during one clinical session. Measures of peak isometric force during resisted abduction and external rotation, and measures of passive ROM abduction and external rotation using the HHD to standardise overpressure force also demonstrated clinically acceptable levels of intraexaminer reliability.

3.2. Interexaminer Reliability of Orthopaedic Special Tests Used in the Assessment of Shoulder Pain

Abstract

Orthopaedic special tests (OST) are commonly used in the assessment of the painful shoulder to assist to rule-in or rule-out specific pathology. A small number of tests with high levels of diagnostic accuracy have been identified but interexaminer reliability data is variable or lacking. The aim of this study was to determine the interexaminer reliability of a group of OST with demonstrated diagnostic accuracy at primary care level. Forty consecutive participants with shoulder pain were recruited. Six tests were performed by two examiners (physiotherapists) on the same day. Tests included the active compression test, Hawkins-Kennedy test, drop-arm test, crank test, Kim test and belly-press test. 'Fair' reliability (kappa 0.36-0.38) was observed for the active compression test (labral pathology), Hawkins-Kennedy test and crank test. Prevalence of positive agreements was low for the active compression test (acromioclavicular joint), drop-arm test, Kim test and belly-press test. Prevalence and bias adjusted kappa (PABAK) values indicated 'substantial' reliability (0.65-0.78) for these tests. The active compression test (acromioclavicular joint), belly-press tests (observation and weakness), Kim test and drop-arm test demonstrate acceptable levels of interexaminer reliability in a group of patients with sub-acute and chronic shoulder conditions.

Introduction

The diagnosis of shoulder pain presents a significant challenge to the primary care clinician due to complex regional anatomy and the frequent coexistence of multiple pathologies. In order to reach a differential diagnosis of shoulder pain, clinicians commonly use orthopaedic special tests (OST) during the physical examination to assist

with ruling-in, or ruling-out specific pathology (Cyriax, 1982). The results of these tests frequently form the basis for diagnostic and intervention decisions.

For a test to be clinically valid, acceptable levels of diagnostic accuracy and interexaminer reliability must be demonstrated (Fritz & Wainner, 2001). Of the OST used in the clinical examination of the painful shoulder, only a small number have demonstrated sufficient diagnostic accuracy to be of clinical use. The Hawkins-Kennedy test has been advocated as a useful screening test for 'impingement' lesions, and the belly press test (subscapularis muscle tear), and active compression test (acromioclavicular joint) are suggested to be specific for their respective pathologies (Hegedus et al., 2008). Anatomical validity has also been established for the Hawkins-Kennedy test and the active compression test (R. Green, Shanley, Taylor, & Perrott, 2008). The belly press test has demonstrated validity for primary activation of the subscapularis muscle (Tokish, Decker, Ellis, Torry, & Hawkins, 2003). The crank test (superior glenoid labrum tears) (Liu et al., 1996; Mimori et al., 1999) and the Kim test (posterior glenoid labrum tears) (S. H. Kim et al., 2005) have demonstrated high levels of diagnostic accuracy for glenoid labral lesions, and the drop-arm test for a complete tear of supraspinatus (Codman, 1934; Murrell & Walton, 2001; Park et al., 2005).

While previous authors provide valuable analyses of the diagnostic accuracy and anatomical validity of orthopaedic special tests, reliability data on many of these tests is lacking, and where available, demonstrates widely variable interexaminer reliability (Table 3.4). In many of these studies, confidence intervals and raw agreement statistics (percent agreement) are not reported, and only one study was identified in which participants were recruited from primary care (Johansson & Ivarson, 2009). No studies investigating interexaminer reliability of the belly press test were found.

The aim of this study was to determine the interexaminer reliability of two experienced physiotherapists in determining the results of a selection of orthopaedic special tests with known diagnostic accuracy used in the assessment of shoulder pain in patients recruited from primary care. The results will inform the content of a clinical examination to be used as index tests in future diagnostic studies, and serve as a guide for the clinician to the selection of evidence-based OST for use in clinical examination of the painful shoulder.

Table 3.4. *Summary of Interexaminer Reliability Studies of Orthopaedic Special Tests for the Shoulder*

Tests	Author	Participant numbers	Examiners	Kappa (95% CI)	% agreement
Impingement tests					
Hawkins-Kennedy test	Johansson & Ivarson (2009)	33	physiotherapists	0.91*	NR
	Nanda et al., (2008)	63	orthopaedic consultant and registrar	0.55*	95%
	Ostor et al., (2004)	136	consultant, specialist registrar & nurse	0.18 - 0.43*	NR
	Razmjou et al., (2004)	136	orthopaedic surgeon & physical therapist	0.29 (0.18, 0.40)	60%
	Norregaard et al., (2002)	86	orthopaedic surgeon & rheumatologist	0.07 - 0.40*	
Glenoid labrum tests					
Kim test	Kim et al., (2005)	172	orthopaedic surgeons	0.91*	NR
Crank test	Walsworth et al., (2008)	55	orthopaedic surgeons & physical therapist	0.20 (-0.05, 0.46)	60%
Active compression test	Walsworth et al., (2008)	55	orthopaedic surgeons & physical therapist	0.24 (-0.02, 0.50)	60%
Rotator cuff integrity					
Drop-arm test	Nanda et al., (2008)	63	orthopaedic consultant and registrar	0.35*	77%
	Ostor et al., (2004)	136	consultant, specialist registrar & nurse	0.28 - 0.66*	NR

Abbreviation: NR, not reported
*confidence interval not reported

Methods

Participants

Consecutive participants were recruited through physiotherapy practices in Christchurch, New Zealand. Participants were included in the study if they were over 18 years of age and currently experiencing shoulder pain. Participants were excluded where pain was referred from a source other than the shoulder, or if there was a history of fracture or dislocation to the shoulder. The study was approved by the Ministry of Health Ethics Committee.

Procedures

Orthopaedic special tests were performed by two experienced examiners (19 and 38 years experience) on the same day. In order to prevent the occurrence of systematic differences between the examiners due to repeated testing and changes in participants' symptom response following the first assessment, the sequence of the examiners was randomly allocated. The order of tests was also randomized using a random sequence generator for each participant and each examiner. Prior to the study, the examiners underwent four training sessions to standardise test procedures and to familiarise themselves with the use of the hand-held dynamometer used to measure peak isometric muscle force during the belly press test.

The orthopaedic special tests were selected according to diagnostic accuracy values and those identified as being of clinical value by Hegedus et al., (2008). The associated criteria for a positive test result are summarized in Table 3.5. The OST were carried out as described by the original authors of the tests (Barth et al., 2006; Codman, 1934; Gerber, Hersche, & Farron, 1996; Hawkins & Kennedy, 1980; S. H. Kim et al., 2005; Liu et al., 1996; O'Brien et al., 1998).

During the belly press test, peak isometric force was recorded and weakness was used as an additional criterion for a positive test result. Peak force was recorded using a hand-held dynamometer (Industrial Research Ltd, Christchurch, New Zealand) stabilized against the participants' abdomen. The device was calibrated on the day of testing to within $\pm$ 0.1kg. Three trials were performed on each arm. The duration of the contraction was approximately 6-7 seconds and trials were followed by approximately 30 seconds rest. Each examiner was blinded to the results of the other and there was no communication between examiners.

Table 3.5. *Response Criteria for Orthopaedic Special Tests*

Test Procedure	Test result	Positive response criteria
Active compression test:		
acromioclavicular joint	+ve/-ve	Pain 'on top' of the shoulder (acromioclavicular joint) that was worse in the position of internal rotation, and relieved or abolished in the position of external rotation/supination.
labral Pathology	+ve/-ve	Pain or a click located 'inside' the shoulder that was worse in the position of internal rotation, and relieved or abolished in the position of external rotation/supination.
Hawkins-Kennedy test	+ve/-ve	Reproduction of participants' symptoms
Drop-arm test	+ve/-ve	An inability to hold the arm at 90 degrees abduction, or a sudden drop of the arm when downward pressure is applied.
Crank test	+ve/-ve	click produced during the test
Kim test	+ve/-ve	production of posterior shoulder pain during the test
Belly-press test:		
observation	+ve/-ve	Patient used shoulder extension to try to exert pressure resulting in elbow dropping behind body.
weakness	+ve/-ve	Weakness of 30% or more compared with the opposite shoulder measured with a hand-held dynamometer (Industrial Research Ltd).

Statistical Methods

Statistical analysis was carried out using Statistical Package for the Social Sciences (SPSS) version 16.0. Percent agreement between examiners, and Cohen's chance-corrected kappa statistics with associated 95% confidence intervals (CI) were calculated for the results of OST. Prevalence and bias adjusted kappa statistics (PABAK) were also calculated to account for unbalanced agreement category scores (prevalence) and differences in proportions of positive and negative results (bias) that are known to adversely affect overall kappa statistics (Feinstein & Cicchetti, 1990; Landis & Koch, 1977; Rigby, 2000; Shankar & Bangdiwala, 2008).

To determine reliability between examiners for the presence of weakness during the belly press test, peak isometric force data was used to calculate a percentage strength deficit of the affected side compared with the unaffected side, then converted to dichotomous values using a 30% or greater deficit as the criteria for a 'positive' response. Data from the peak force trial, and mean of three trials was used in the analysis. Two by two contingency tables were constructed for the results of the two examiners, and kappa statistics with associated 95% confidence intervals, PABAK and percent agreement statistics were calculated.

To determine whether extremes of prevalence or bias were likely to affect the overall kappa value, the prevalence index[1] (PI) and bias index[2] (BI) were calculated for each variable according to Byrt (1993). Prevalence index values can range from -1 to +1, and the prevalence index is equal to zero when 'yes' and 'no' are equally probable (Byrt, Bishop, & Carlin, 1993). Bias index values can range from zero to 1, and equal zero only if there is no difference in 'positive' proportions between examiners (Byrt et al., 1993).

Where the prevalence index was high, the PABAK value was used for interpretation of results. For the purposes of this study, an arbitrary cut-off value of a prevalence index less than -0.5, or greater than 0.5, was used for interpretation of the PABAK values instead of overall kappa scores. Kappa and PABAK values for interexaminer reliability were interpreted according to the guidelines of Landis and Koch (1977); <0.00 poor; 0.00-0.20 slight; 0.21-0.40 fair; 0.41-0.60 moderate; 0.61-0.80 substantial; 0.81-1.00 almost perfect (Landis & Koch, 1977).

Results

Forty participants with shoulder pain were recruited. Participants included 23 males and 17 females with a mean age of 49 years (range 18 – 77 years). Descriptive data for participants are summarized in Table 3.6. Randomization of examiner order resulted in 21 participants being examined first by examiner 1 and 19 by examiner 2. Nine participants with bilateral shoulder pain were excluded from the analysis of 'weakness' during the belly press test.

Overall kappa values, raw prevalence of positive test results, prevalence index, PABAK and percent agreement statistics for results of OST are presented in Table 3.7. Bias index results (not presented) ranged from 0.03 – 0.23 indicating the PABAK values were not significantly affected by examiner bias. On this basis, the PABAK values are interpreted as predominantly reflecting prevalence adjustment.

Interexaminer agreement on the results of OST was varied. Overall kappa values ranged from -0.04 to 0.65 ('poor' to 'substantial'). The prevalence of positive test results ranged from 15% (active compression test -acromioclavicular joint and Kim test) to 75% (active compression test -labral pathology). The prevalence index exceeded -0.50 and 0.50 for a number of tests (active compression test -acromioclavicular joint,

---

[1] Prevalence Index = total number of 'positive' – total number of 'negative agreements / number of cases.
[2] Bias Index = Difference in proportions of 'yes' for the two examiners / number of cases.

drop-arm test, Kim test and belly press test – observation and weakness). The PABAK values ranged from 0.65-0.78 for indicating 'substantial' agreement between examiners for these tests. Percent agreement ranged from 83% to 89%.

Table 3.6. *Summary of Participant Characteristics*

Characteristics		Number	%
Gender	male	23	58
	female	17	42
Affected Side	dominant	24	60
	non-dominant	11	28
	bilateral	5	13
		Mean (median)	**Range**
Age (years)		49 (51)	18-77
Height (cm)		171 (169)	157-189
Weight (kg)		80 (84)	53-102
Duration of symptoms (months)		48 (8)	<1-325
Pain severity in previous 24 hours (11 point VAS)		3.6 (4.0)	0-7

Abbreviation: VAS, visual analogue scale

Tests where prevalence was not considered to adversely affect the kappa value included the active compression test (labral pathology), Hawkins-Kennedy test and crank test. 'Fair' agreement was demonstrated for these tests (kappa 0.36 -0.38), and confidence intervals for the overall kappa values were wide. Percent agreement values for these tests ranged from 68% to 70%.

Highest levels of interexaminer agreement were observed for the belly press (weakness) (PABAK 0.78) and lowest agreement was observed for the crank test (kappa 0.36).

Table 3.7. *Interexaminer Reliability of Orthopaedic Special Tests*

Orthopaedic special test	Positive results (% of total cases)	Prevalence index[a]	Overall kappa (95% CI)	PABAK[b]	% agreement
Active compression test:					
acromioclavicular joint	6 (15%)	0.83	0.22 (-0.24, 0.68)	0.75	88
labral pathology	30 (75%)	-0.25	0.38 (0.1, 0.65)	0.40	70
Hawkins-Kennedy test	27 (68%)	-0.03	0.38 (0.10, 0.63)	0.35	68
Drop-arm test	7 (18%)	-0.78	0.57 (-0.14, 0.57)	0.67	83
Crank test	18 (45%)	-0.35	0.36 (-0.07, 0.59)	0.35	68
Kim test	6 (15%)	-0.85	-0.04 (-0.12, 0.03)	0.70	85
Belly press test:					
observation	9 (23%)	-0.73	0.31 (-0.03, 0.64)	0.65	83
weakness (maximal trial)	10 (25%)	-0.61	0.65 (0.33, 0.96)	0.78	89
weakness (mean of 3 trials)	11 (28%)	-0.58	0.58 (0.26, 0.90)	0.72	86

Abbreviations. CI, confidence interval; PABAK, prevalence and bias adjusted kappa

[a]Prevalence Index = total number of 'positive' – total number of 'negative agreements / number of cases (a-d/N)

[b]PABAK = $\dfrac{(2n/N)\text{-}0.5}{1\text{-}0.5}$ = $2p_o$-1

Discussion

Kappa or PABAK values in excess of 0.60, and percent agreement in excess of 80% are required for a test to be considered appropriate for inclusion in a clinical examination. The active compression test (acromioclavicular joint), drop-arm test, Kim test and belly press test (observation and weakness), all reached this level in the present study. Interexaminer reliability results for the drop-arm test, crank test in the current study are similar to those obtained where examiners were trained in orthopaedics and had a special interest in shoulders (Nanda, Gupta, Kanapathipillai, Liow, & Rangan, 2008; Norregaard et al., 2002; Ostor, Richards, Prevost, Hazleman, & Speed, 2004; Walsworth et al., 2008). In the present study, the prevalence of positive results for the drop-arm test was low, and further studies using larger numbers are required to confirm this result in the primary care environment.

Interexaminer reliability of the Hawkins-Kennedy test has been previously reported as 'fair' between an orthopaedic surgeon and a physical therapist (kappa 0.29; 95% CI 0.18, 0.40) (Razmjou, Holtby, & Myhr, 2004). The only previous study identified involving symptomatic primary care patients in which both examiners were physiotherapists reported considerably higher reliability between examiners than observed in the current study (kappa 0.91; CI not reported) (Johansson & Ivarson, 2009). These authors investigated only four tests to identify subacromial pain. Differences between these results may be partially explained by the higher number of tests conducted in the present study, and the resulting potential for random error as a result of a change in the participants' symptoms between assessments.

The results of the Kim test in the current study (PABAK 0.70) differ from those of the original authors who reported an ICC value of 0.91 (S. H. Kim et al., 2005). However, the original study was conducted by the physician who developed the test and the methods and procedures for collection of interexaminer reliability data were not described. The Kim test demonstrated 'substantial' interexaminer reliability according to prevalence adjusted statistics, and high levels of agreement (85%) in the present study. Verification in a larger sample of the primary care population is required.

The results of this study provide the first known interexaminer reliability data for the belly press test. Both components of the belly press test (observation and weakness) demonstrated clinically acceptable levels of interexaminer reliability according to prevalence adjusted statistics (PABAK 0.65-0.78), and amongst the highest levels of

raw agreement (83-89%) suggesting this is a reliable method of assessing the integrity of subscapularis.

The active compression test is reported to differentiate between acromioclavicular joint and glenoid labrum pathology (O'Brien et al., 1998). Previous studies indicate 'fair' reliability (kappa 0.24; 95%CI -0.02, 50) for the combined result of the active compression test for both pathologies (Walsworth et al., 2008). No studies were identified that tested the interexaminer reliability of differentiated test results of the active compression test for the two separate pathologies. Results indicated a higher level of raw agreement between examiners (88%) and prevalence adjusted reliability (PABAK 0.75) in determining a positive result for acromioclavicular joint pain compared with labral pathology (raw agreement 70%; kappa 0.38). This finding may be explained by the relative ease with which participants identify the more definitive, superficial pain localized to the "top" of the shoulder (acromioclavicular joint pain) compared with non-specific pain "inside" the shoulder.

Reported high pain severity and longstanding complaints have previously been identified as determinants of disagreement in diagnostic classification studies (de Winter et al., 1999). In this study, the mean duration of symptoms was high (48 months), which is longer than the duration of symptoms typically reported by patients in the primary care setting. Therefore the results of this study may not represent interexaminer reliability of these tests in patients with shoulder pain of shorter duration.

Conclusion

The active compression test (acromioclavicular joint pathology), drop-arm test, Kim test and belly press test demonstrated acceptable levels of interexaminer reliability when corrected for the low prevalence of positive results. Test reliability is a prerequisite for diagnostic validity, and the false positive and false negative rate of these and other tests will be evaluated to estimate their diagnostic accuracy in the primary care population. Results will be reported in due course.

3.3. RELIABILITY OF SYMPTOM RESPONSES ASSOCIATED WITH RANGE OF MOTION AND RESISTED TESTS

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Williams, M. Reliability of symptom responses associated with shoulder range of motion and resisted muscle tests.

The results presented in the following manuscript have been prepared for journal submission upon completion of the PhD.

Abstract

Symptom reproduction during physical examination tests may inform diagnostic reasoning with respect to involvement of specific tissues in pain generation, however, reliability of this clinical variable has received little attention. The aim of this study was to evaluate the interexaminer reliability of symptom responses associated with range of motion and resisted muscle tests used in the assessment of shoulder pain. Forty patients with shoulder pain were consecutively recruited from primary care and examined by two examiners on the same day. Examiners recorded the presence of a painful arc during abduction, and symptom responses during active and passive ROM tests and resisted tests of the shoulder. Agreement and overall kappa values were highest for symptom responses during painful arc abduction (77% agreement; kappa 0.66) and active ROM flexion (80% agreement; kappa 0.60). Passive ROM cross-body adduction (external rotation) demonstrated the lowest levels of agreement (61%; kappa 0.22), and 'moderate' agreement was observed for all other tests (agreement 70% to 76%; kappa 0.42 to 0.47). Symptom responses associated with active ROM tests were consistently reported by two examiners and may aid diagnostic differentiation of inert or contractile structures in those with shoulder pain. Symptom responses associated with passive ROM and resisted tests may also be of value.

Introduction

Primary care practitioners, including physiotherapists and general medical practitioners, commonly use a method of clinical examination of the shoulder girdle that is based upon the method of orthopaedic assessment described by Cyriax (Cyriax, 1982). This method of assessment aims to reproduce patient symptoms using tests designed to differentiate between pain produced by inert (capsule-ligamentous) or

contractile structures using ROM and resisted muscle tests (Cyriax, 1982). Reproduction of the patient's symptoms during ROM and resisted muscle tests thus represents an important clinical feature, assisting the clinician to identify the specific tissue involved in pain production. Despite the growing body of knowledge surrounding reliability of measurement of shoulder ROM, little evidence is available regarding the interexaminer reliability of patient symptom responses.

Pain responses ranked on an ordinal scale (no pain − excruciating pain) demonstrated 'fair' to 'substantial' agreement between examiners (Landis & Koch, 1977) for a range of physical examination tests of the shoulder (kappa 0.35-0.69; agreement 73%-91%) (K. Hayes & Petersen, 2001). Variable agreement between examiners has also been reported regarding the sequence of onset of pain and resistance during shoulder ROM tests (kappa 0.13 − 0.62) (K. Hayes & Petersen, 2001). No studies were identified that rated interexaminer agreement on the presence or absence of symptoms (dichotomous response).

By using resisted tests for specific muscles the clinician aims to differentiate which contractile structure may be involved in pain production according to the patient's symptom response and degree of weakness (Cyriax, 1982). However, agreement between two experienced examiners on symptom response during resisted tests has been found to be only "slight" to "fair" (kappa statistic <0.40), with only shoulder abduction and elbow extension producing "moderate" levels of agreement (kappa >0.40) (K. W. Hayes & Petersen, 2003).

The aim of the study was to evaluate the interexaminer reliability of symptom responses associated with range of motion and resisted muscle tests used in assessment of shoulder pain in a symptomatic population of patients recruited from primary care. The results will inform the content of a clinical examination to be used as the index test in the diagnostic study.

Methods

Design, sampling, recruitment and test procedures have been described in the previous sections. These consisted of recording symptom responses associated with painful arc abduction, active ROM elevation (flexion) and hand-behind-back, passive ROM glenohumeral abduction, external rotation and cross-body adduction, and resisted abduction and external rotation. Symptom responses were recorded as "positive" if the

test procedure reproduced the participants' typical pain, regardless of pain intensity. Responses were recorded as negative if no pain or 'unfamiliar' pain was produced.

Statistical Methods

Percentage agreement between examiners, and Cohen's chance-corrected kappa statistics with associated 95% confidence intervals were calculated for the presence of painful arc of abduction and symptom response (SR) associated with ROM and resisted muscle tests. To determine whether extremes of prevalence or bias affected the overall kappa value, the prevalence index[3] (PI) and bias index[4] (BI) were calculated for each variable according to Byrt (1993). For the purposes of this study, an arbitrary cut-off value of a prevalence index less than -0.5, or greater than 0.5 was selected for interpretation of the PABAK values instead of overall kappa scores.

Results

Forty participants with shoulder pain were recruited. Participants included 23 males and 17 females with a mean age of 49 years (range $18 - 77$ years). Descriptive data for participants have been previously summarized (Table 3.6). Prevalence index values ranged between 0.03 and 0.34 hence kappa values were used for interpretation of results (Table 3.8).

Symptom responses during painful arc abduction and active ROM flexion demonstrated highest levels of agreement (77% to 80%), with kappa values of 0.66 and 0.60 respectively (Table 3.8). Passive ROM cross-body adduction (external rotation) demonstrated the lowest levels of agreement (61%; kappa 0.22), and 'moderate' agreement was observed for all other tests (agreement 70% to 76%; kappa 0.42 to 0.47) (Table 3.8).

Discussion

The aim of this study was to investigate the interexaminer reliability of recorded symptom responses during physical examination tests of the shoulder. Despite this being a variable upon which diagnostic decisions are frequently made, the reliability of this clinical variable has received little attention to date.

---

[3] Prevalence Index = total number of 'positive' – total number of 'negative' agreements / number of cases.

[4] Bias Index = Difference in proportions of 'yes' for the two raters / number of cases.

Table 3.8. *Interexaminer Reliability for Symptom Responses Associated with Physical Examination Tests*

Test	% agreement	Prevalence index	Kappa (95% CI)
Painful arc abduction	77	0.14	0.66 (0.27-0.81)
Active ROM flexion	80	0.10	0.60 (0.35-0.85)
Active ROM hand behind back	73	0.03	0.45 (0.18-0.73)
Passive ROM abduction	70	0.10	0.42 (0.16-0.67)
Passive ROM external rotation	73	0.08	0.47 (0.23-0.71)
Passive ROM cross-body adduction (internal rotation)	76	0.34	0.47 (0.17-0.76)
Passive ROM cross-body adduction (external rotation)	61	0.03	0.22 (-0.07-0.52)
Resisted abduction	75	0.25	0.47 (0.19-0.75)
Resisted external rotation	71	0.08	0.43 (0.15-0.70)

Abbreviations. CI, confidence interval; ROM, range of motion.

Overall, agreement was high for symptom responses during active ROM tests, compared with symptom responses associated with passive ROM and resisted muscle tests. Agreement between examiners regarding the presence or absence of a painful arc of motion during abduction was 'substantial' (kappa 0.66) and is higher than that reported by other investigators (Nanda et al., 2008; Norregaard et al., 2002; Ostor et al., 2004). Of the active ROM tests, less agreement was reported between examiners for symptom responses associated with hand-behind-back. This is a complex movement involving components of humeral extension, adduction and internal rotation, scapula rotation and elbow flexion (Ginn, Cohen, & Herbert, 2006; Mallon, Herring, Sallay, Moorman, & Crim, 1996). Despite clear instructions to participants to provide a 'positive' response only if their 'typical' symptoms were produced, some participants may have confused generalized discomfort during this movement with their symptoms.

Agreement regarding symptom responses associated with passive ROM tests and resisted tests, although 'moderate', reflected less agreement than those recorded for active ROM tests. Variations between examiners in the magnitude of applied force at end during resisted muscle tests may have affected the number of positive responses, with larger magnitudes of force more likely to provoke symptoms. Previous authors noted that the majority of disagreements on diagnostic classification were related to pain response with resisted tests (K. W. Hayes & Petersen, 2003). Force overpressure was standardised during measures of passive ROM, however interpretation by the subject of end-range discomfort as representing typical symptoms may have affected the observed agreement levels for these tests. Standardisation of force and resistance during passive ROM and resisted muscle tests, and clear differentiation between discomfort and typical

symptom reproduction would appear to be important factors for interexaminer reliability of these tests.

Conclusion

Symptom responses associated with active ROM tests of the shoulder in patients with predominantly sub-acute and chronic shoulder pain are reliable between experienced examiners. Symptom responses associated with passive ROM and resisted tests of the shoulder demonstrate moderate agreement between examiners. Further standardisation of test procedures may yield improvements in agreement for this clinical feature.

3.4. Methodological Development for Measures of Range of Motion and Peak Muscle Force.

This section contains unpublished results that were used to inform the procedures and content of the diagnostic accuracy study.

Abstract

Measurements of ROM and strength (peak isometric force) using a new hand-held dynamometer demonstrated moderate measurement variability in a previous study conducted as part of this thesis. The aim of this pilot study was to re-evaluate intra-examiner reliability following further standardisation of ROM and resisted test procedures. A chair was manufactured in order to standardise the starting position for measures of passive ROM and resisted tests. External stabilisation attachments were used for measures of peak muscle force during resisted tests. Twelve volunteers were recruited and examined by one examiner in one session. Tests consisted of three trials of active elevation (flexion) and passive ROM (abduction, external and internal rotation) and resisted muscle tests (abduction, external rotation and internal rotation). For ROM tests (single trials), measures of active ROM elevation demonstrated the lowest measurement variability (95% LOA 4.2^o) and passive ROM internal rotation demonstrated the highest variability (95% LOA 14.3^o). For the mean of three trials all LOA were less than 10°. For resisted tests, 95% LOA ranged from 0.7kg (resisted external rotation (unaffected side) to 2.4kg (resisted internal rotation (unaffected side). Limits of agreement were less than 2.5kg, and ICCs exceeded 0.90 for both single trials, and multiple trials for all tests. These results demonstrate reduced measurement variability following standardisation of test procedures compared with previous results. Single trials of active ROM elevation, passive ROM abduction, passive ROM external rotation (0° abduction) and resisted tests demonstrated acceptable intraexaminer measurement reliability. However three trials of passive ROM external and internal rotation (90° abduction) are required for acceptable measurement reliability.

Introduction

Measurements of shoulder range of motion (ROM) and strength are commonly used in the diagnosis and to assess the progress of treatment in patients suffering from shoulder pain. Agreement between examiners on measures of ROM and strength is an

important component of diagnostic test validity to ensure consistent diagnostic classification when tests are applied in a wide range of settings by a number of examiners. Measurement accuracy and consistency within individual practitioners is also important to accurately assess clinical progress and the effects of treatment interventions.

Results of a previous study indicated considerable measurement variability between examiners for measures of shoulder passive range of motion (ROM) (ICC 0.45 – 0.58) and measures of peak isometric muscle force (0.68 – 0.84) (Cadogan, Laslett, et al., 2011a). Moderate to high levels of absolute measurement variability were also demonstrated within examiners for measures of ROM (6-19°) and peak isometric muscle force (1.1-7.0kg).

Although measures of active ROM elevation demonstrated high levels of relative reliability in the previous study (ICC 0.88-0.99), absolute differences (limits of agreement (LOA)) within- (6-19°) and between examiners (14-22°) demonstrated moderate measurement variability. This test was performed with the participant in the standing position. It is likely that compensatory postural movements, including lumbar extension and changes in symptoms and ROM as a result of repeated testing, accounted for some of the observed measurement variability.

Measures of passive ROM have previously demonstrated high levels of variability, frequently reported to be due to the inability to standardise the amount of overpressure applied at end range of motion (Boone et al., 1978; Gajdosik & Bohannon, 1987). However in a previous study (Cadogan, Laslett, et al., 2011a), little improvement in interexaminer reliability was observed with standardisation of force overpressure at end range of motion using a hand-held dynamometer. A lack of familiarisation with the hand-held dynamometer (HHD) and insufficient standardisation of test procedures are among possible explanations for this result.

The measurement reliability of peak isometric muscle force using a new HHD was evaluated in previous work in which the aim was to gather baseline information on measurement variability for the new device (Cadogan, Laslett, et al., 2011a). Results indicated that measures of peak isometric abduction, external rotation and internal rotation using a HHD demonstrated only moderate interexaminer reliability (ICC 0.68-0.84), and wide intraexaminer limits of agreement (LOA) (1.1-7.0kg). Factors reported to influence measurement variability when using a HHD include examiner gender, weight, strength, variation in testing protocol and lack of stabilisation of the measuring

device (Agre et al., 1987; Bohannon, 1999; Wadsworth et al., 1992). Others have reported improvements in measurement repeatability for shoulder strength using stabilisation of the HHD (ICC 0.97; SEM 0.62-1.15kg) (Kolber, Beekhuizen, Cheng, & Fiebert, 2007). In the previous study examiners differed with respect to gender and weight, and the HHD was not stabilised during testing. When diagnostic criteria depend upon accurate measures of strength and associated deficits for conditions such as rotator cuff tears (McCabe et al., 2005), accurate and repeatable measures of strength are required to avoid incorrect diagnostic classification.

Measures of passive ROM shoulder external and internal rotation performed at 90° of abduction, and resisted internal rotation are also relevant clinical variables for the diagnosis of specific shoulder pathology such as adhesive capsulitis, internal impingement or lesions affecting the subscapularis component of the rotator cuff (Gerber et al., 1996; McFarland, Hsu, Neira, & O'Neil, 1999; Zuckerman, Cuomo, & Rokito, 1994). These variables were not included in the original reliability study and evaluation of intraexaminer measurement reliability was required prior to inclusion of these clinical tests in the diagnostic accuracy study.

The aim of this study was to assess the intraexaminer reliability of measures of shoulder ROM and peak isometric force during resisted shoulder tests using strictly controlled test procedures and external stabilisation of the HHD.

Methods

Participants

A sample of convenience was recruited from QEII Stadium, Christchurch, New Zealand. Staff working in this complex were contacted via email and asked to volunteer for this study if they were currently suffering from shoulder pain.

Hand-Held Dynamometer

The hand-held dynamometer (Industrial Research Ltd, Christchurch, New Zealand) (Cadogan, Laslett, et al., 2011a) was used to measure passive ROM (degrees) and peak isometric muscle force (kg) during selected tests. Standardisation of force-overpressure at end range of motion during passive glenohumeral abduction (6% of body weight) and external rotation (3kg) was used as previously described (Cadogan, Laslett, et al., 2011a).

Procedures

Prior to testing, the examiner underwent extensive familiarization with the HHD (approximate total familiarization time 12 hours). In order to overcome the postural and compensatory movements observed in participants during ROM and resisted tests in the initial reliability study, an aluminium-framed chair was developed and constructed by Industrial Research Ltd (Christchurch, New Zealand) (Figure 3.2)**.** The chair consisted of several features designed to standardise test position and to eliminate compensatory movements. Features included an angled seat designed to help prevent lumbar extension during arm elevation, an adjustable-height thoracic backboard to help standardise upright sitting posture during resisted tests, adjustable shoulder stabilisers to help reduce shoulder elevation during measures of passive glenohumeral abduction and compensatory shoulder elevation during resisted tests, and attachment sites for external stabilisers for the HHD.
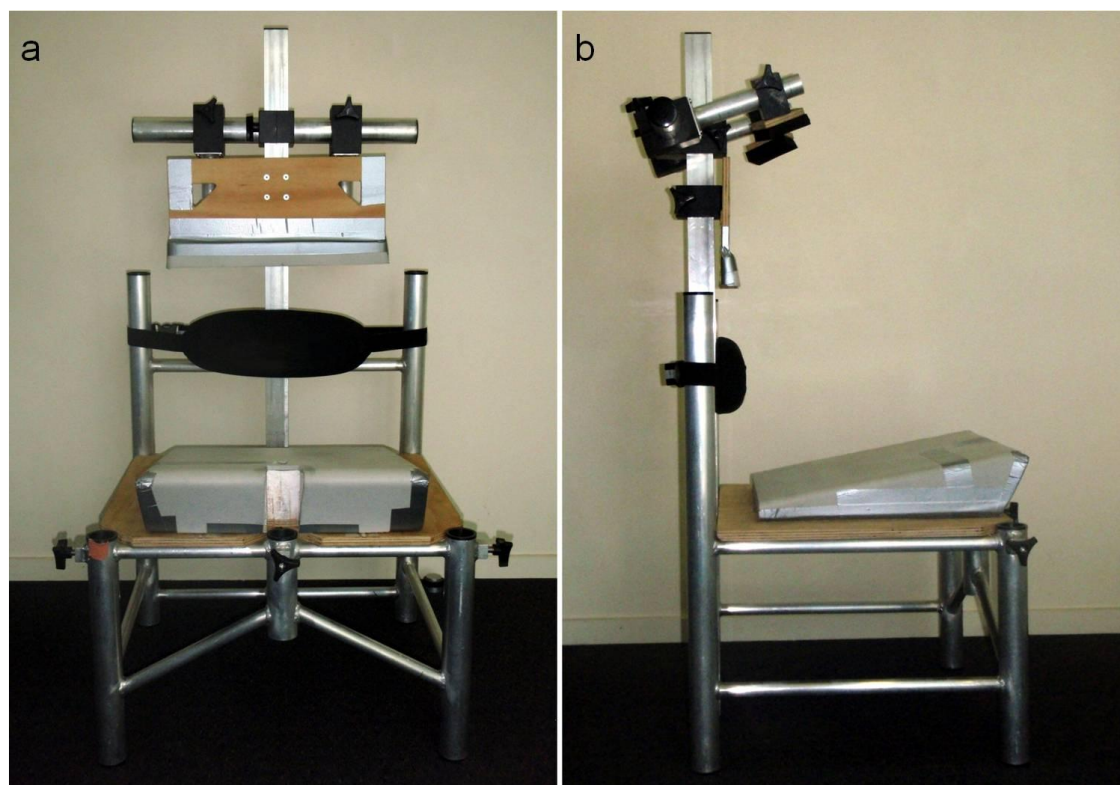
Figure 3.2. Chair developed for standardisation of test procedures. Figure shows chair from (a) front view; b) side view).

Using the standardised procedures (Table 3.9) the following measures were performed three times for each participant by a single examiner (AC):

▪ Active ROM elevation (flexion) (Figure 3.3)

- Passive ROM glenohumeral abduction, external rotation (0° and 90° abduction), internal rotation (90° abduction) (Figure 3.4 a-d)
- Peak isometric force during resisted abduction, external rotation and internal rotation (using external stabilization of HHD) (Figure 3.5 a-c).
- Peak isometric internal rotation force during the Belly press test (Figure 3.5 d).

Statistical Methods

Relative reliability was assessed using single-measure intraclass correlation coefficients ($ICC_{2,1}$) and associated 95% confidence intervals (CI) (two-way random effects model -absolute agreement). Data from the first trial was compared with data from the second trial, and the mean of three trials for ROM and peak muscle force measures. A one-way ANOVA was used to ascertain any differences between trials with the level of significance set at $p=0.05$.

Absolute reliability was determined by calculating the mean difference between measures and the associated 95% CI for the mean difference, as well as limits of agreement (LOA) according to the Bland and Altman method of assessing agreement (mean difference between trials $\pm$ 1.96 SD_{diff}) (Bland & Altman, 1986).

Reliability values were interpreted according to the guidelines of Landis and Koch (1977); 0.00-0.20 slight; 0.21-0.40 fair; 0.41-0.60 moderate; 0.61-0.80 substantial; 0.81-1.00 almost perfect (Landis & Koch, 1977).

Table 3.9. *Description of Operational Test Procedures Following Methodological Development*

Test	Participant position and instructions	Criterion for end of test
Active ROM		
elevation (flexion)	Participant seated in chair with lumbar spine in contact with lumbar support. Participant was instructed to raise both arms to the ceiling keeping lumbar spine in contact with support, elbows straight and hands shoulder width apart.	When prevented from going further by pain or movement limitation, or when lumbar spine lost contact with support.
Passive ROM		
glenohumeral abduction	Participant seated in chair. Attachment set up along axis of glenohumeral joint and board attached. Shoulder pads positioned to prevent shoulder girdle elevation. Participant instructed to relax the arm.	HHD alarm sounded (6% body weight) or when participants said "stop" (maximal tolerable pain limit reached).
external rotation (0° abduction)	Participant supine, knees extended, pad under elbow to maintain arm in plane of scapula. Participant instructed to relax as arm passively moved into external rotation.	HHD alarm sounded (3kg), when participants said "stop" (maximal tolerable pain limit reached), or compensatory movement occurred.
external rotation (90° abduction)	Participant supine, knees bent, pad under elbow to maintain arm in plane of scapula. Participant instructed to relax.	HHD alarm sounded (3kg), when participants said "stop" (maximal tolerable pain limit reached), or compensatory movement occurred.
internal rotation (90° abduction)	Participant supine, knees bent, pad under elbow to maintain arm in plane of scapula. Participant instructed to relax as arm passively moved into internal rotation.	HHD alarm sounded (3kg), when participants said "stop" (maximal tolerable pain limit reached), or compensatory movement occurred.
Resisted tests		
abduction	Participant seated in chair. Elbow flexed to 90° and shoulder abducted to approximately 10°, back of head against chair. Attachments adjusted for participant limb position. The HHD was mounted onto platform with force pad against lateral humerus immediately proximal to lateral epicondyle. Participant instructed to build up force to maximum levels, maintain until examiner said "stop" (approx 6 seconds), maintaining head against chair and forearm parallel with attachment arm (to prevent internal rotation).	After approximately 6 seconds contraction time.

external rotation	Participant seated in chair. Elbow flexed to 90°, shoulder in neutral position (0° abduction, 10 to 15° internal rotation), forearm mid-pronation (thumb up) back of head against chair, low friction pad between elbow and thorax to maintain adduction and reduce deltoid contribution. The HHD was mounted onto platform with force pad against the dorsal surface of the distal radius immediately proximal to wrist crease. Participant instructed to build up force to maximum levels, maintaining head against chair and to keep low friction pad in place using adduction. Participant instructed to maintain contraction until examiner said "stop" (approx 6 seconds).	After approximately 6 seconds contraction time or when pad fell from between elbow and thorax.
internal rotation	Participant seated in chair. Elbow flexed to 90°, shoulder in neutral position (0° abduction, 10 to 15° internal rotation), forearm mid-pronation (thumb up) back of head against chair. The HHD was mounted onto platform with force pad against the ventral surface of the distal radius immediately proximal to wrist crease. Participant instructed to build up force to maximum levels, maintaining head against chair and to keep arm in contact with thorax (to prevent abduction). Participant instructed to maintain contraction until examiner said "stop" (approx 6 seconds).	After approximately 6 seconds contraction time or when arm lifted away from side.
belly press test	Participant seated in chair with lumbar spine in contact with support and head on back of chair. No attachments used. Participant instructed to place ventral surface of wrist in middle of abdomen with forearm parallel to ground. The HHD was placed on firm platform on abdomen at level of participants ventral wrist crease. Participant instructed to press into HHD maintaining elbow in front of body, lumbar spine on support and back of head on chair.	After approximately 6 seconds, or when elbow fell behind body, or shoulder rolled forward of body.

Abbreviations. ROM, range of motion; HHD, hand-held dynamometer
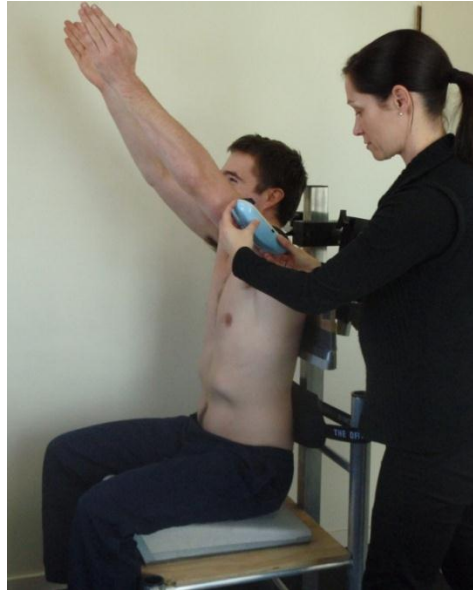
Figure 3.3. Measurement of active range of motion elevation (flexion). A standardised procedure was used to minimise compensatory trunk motion using lumbar and thoracic support.

Figure 3.4. Measurement of passive range of motion. Standardised positions were used to measure: a) glenohumeral abduction; b) external rotation (at 0° abduction); c) external rotation (at 90° abduction); d) internal rotation (at 90° abduction).

Figure 3.5. Measurement of peak isometric force. Standardised positions and external stabilization of the hand-held dynamometer were used to measure peak force during resisted tests: a) abduction; b) external rotation; c) internal rotation; d) internal rotation in the belly-press test position.

Results

Twelve volunteers were assessed. Demographic characteristics are presented in Table 3.10.

Table 3.10. *Participant Demographics*

Characteristics	Median (IQ range)	Range
Age (years)	34 (17)	22-68
Weight (kg)	77 (14)	60-95
Symptom duration (weeks)	11 (11)	3-52
	Number	%
Gender male	10	83
Affected side right	8	67
Dominant side affected	7	58

Abbreviations. IQ range; interquartile range

Range of Motion Tests

The ICCs for all ROM measures on both the affected and unaffected side ranged from 0.76 (passive ROM internal rotation, affected side), to 0.99 (passive ROM external rotation (0° abduction) (Table 3.11). Mean differences between trials ranged from 0.0^o (passive ROM internal rotation) to 7.1^o (passive ROM external rotation (90° abduction), and LOA ranged from 4.2^o (active ROM elevation) to 14.3^o (passive ROM internal rotation). For single trials, LOA were less than 10° for all tests except passive ROM glenohumeral abduction and internal rotation. For the mean of three trials compared with a single trial, LOA were less than 10° and ICCs exceeded 0.80 for all ROM tests.

Resisted Tests

The ICCs for resisted tests ranged from 0.91 (resisted external rotation (unaffected side), to 0.99 (all resisted tests on the affected side)). Mean differences between trials ranged from 0.0kg (resisted abduction (affected side) and resisted external rotation (unaffected side) to 1.0kg (resisted internal rotation (affected side)). Limits of agreement ranged from 0.7kg (resisted external rotation (unaffected side) to 2.4kg (resisted internal rotation (unaffected side). Limits of agreement were less than 2.5kg, and ICCs exceeded 0.90 for comparisons between single trials, and between a single trial and mean of three trials for all tests.

For the belly press test, ICCs ranged from 0.92 to 0.98, mean difference between trials ranged from 0.2 to 0.5 kg and LOA ranged from 1.6kg (unaffected side) to 2.9kg (affected side). Slight reductions in measurement variability and improvements in consistency were seen for the mean of three trials compared with comparisons between single trials.

No consistent trend was observed with regards to mean differences and LOA for the affected compared with the unaffected side for ROM and resisted tests, however ICCs were lower for measures of passive ROM on the affected side (0.76 to 0.96) compared with the unaffected side (0.93 to 0.99). Overall, mean differences between trials and LOA (measurement variability) were smaller, and ICCs (consistency) were higher for the mean of three trials compared with comparisons between single trials.

Table 3.11. *Intraexaminer Results for Measures of Range of Motion and Peak Isometric Force Following Methodological Development*

Tests	No. of trials	Unaffected side			Affected side		
		Mean (range)	Mean diff $\pm$ LOA (95% CI)	ICC (95% CI)	Mean (range)	Mean diff $\pm$ LOA (95% CI)	ICC (95% CI)
Active ROM							
elevation	single trial	171 (154, 178)	0.8 $\pm$ 7.1 (-1.5, 3.0)	0.90 (0.69, 0.97)	171 (159, 181)	0.4 $\pm$ 6.3 (-2.5, 1.6)	0.86 (0.58, 0.96)
	three trials	171 (147, 178)	0.2 $\pm$ 5.0 (-1.5, 1.8)	0.94 (0.81, 0.98)	171 (158, 177)	0.1 $\pm$ 4.2 (-1.5, 1.3)	0.93 (0.79, 0.98)
Passive ROM (degrees)							
glenohumeral abduction	single trial	93 (71, 111)	0.8 $\pm$ 8.1 (-1.5, 3.2)	0.93 (0.79, 0.98)	90 (74, 111)	0.8 $\pm$ 11.0 (-2.8, 4.3)	0.87 (0.61, 0.96)
	three trials	94 (76, 113)	0.9 $\pm$ 4.8 (-0.7, 2.5)	0.97 (0.91, 0.99)	91 (67, 108)	0.5 $\pm$ 6.6 (1.6, 2.6)	0.95 (0.83, 0.98)
external rotation (0° abd)	single trial	63 (31, 82)	2.4 $\pm$ 9.9 (-5.4, 0.6)	0.94 (0.80, 0.98)	63 (49, 84)	0.8 $\pm$ 9.8 (-2.1, 3.8)	0.93 (0.77, 0.98)
	three trials	62 (31, 82)	1.4 $\pm$ 4.7 (-3.0, 0.3)	0.99 (0.95, 1.00)	64 (53, 87)	0.6 $\pm$ 6.5 (-1.6, 2.9)	0.96 (0.85, 0.99)
external rotation (90° abd)	single trial	113 (54, 136)	1.4 $\pm$ 10.8 (-5.1, 2.3)	0.97 (0.88, 0.99)	126 (116, 134)	7.1 $\pm$ 9.6 (-10.4, -3.8)	0.78 (0.37, 0.94)***
	three trials	113 (59, 134)	0.1 $\pm$ 7.8 (-2.6, 2.8)	0.98 (0.93, 1.00)	121 (110, 135)	4.8 $\pm$ 7.2 (-7.3, -2.3)	0.86 (0.56, 0.96)**
internal rotation (90° abd)	single trial	75 (47, 106)	0.6 $\pm$ 11.5 (-3.3, 4.6)	0.91 (0.69, 0.97)	69 (51, 87)	3.8 $\pm$ 14.3 (-1.1, 8.7)	0.76 (0.33, 0.93)
	three trials	75 (47, 98)	0.0 $\pm$ 7.3 (-2.5, 2.5)	0.96 (0.86, 0.99)	71 (57, 88)	2.7 $\pm$ 7.7 (0.1, 5.4)	0.94 (0.79, 0.98)
Peak isometric force (kg)							
abduction	single trial	15.2 (10.2, 23.7)	0.0 $\pm$ 2.3 (-0.8, 0.8)	0.95 (0.84, 0.99)	16.3 (10.4, 32.1)	0.2 $\pm$ 2.1 (-0.5, 0.9)	0.98 (0.94, 0.99)
	three trials	15.2 (10.0, 22.8)	0.0 $\pm$ 1.4 (-0.4, 0.5)	0.98 (0.95, 1.00)	16.4 (10.4, 30.5)	0.1 $\pm$ 1.5 (-0.4, 0.6)	0.99 (0.97, 1.00)

external rotation	single trial	12.0 (8.7, 16.6)	0.6 ± 1.8 (0.0, 1.2)	0.91 (0.71, 0.97)*	11.6 (7.1, 18.5)	0.1 ± 1.7 (-0.6, 0.4)	0.95 (0.84, 0.99)
	three trials	12.2 (9.0, 16.7)	0.3 ± 0.7 (0.0, 0.5)	0.99 (0.95, 1.00)*	11.7 (7.7, 17.8)	0.0 ± 0.9 (-0.3, 0.3)	0.99 (0.96, 1.00)
internal rotation	single trial	14.5 (10.4, 24.6)	1.0 ± 2.4 (0.2, 1.8)	0.95 (0.83, 0.99)*	14.3 (10.4, 28.1)	0.9 ± 1.7 (0.3, 1.5)	0.98 (0.94, 1.00)**
	three trials	15.2 (11.8, 26.3)	0.7 ± 1.7 (0.2, 1.3)	0.98 (0.91, 0.99)	15.9 (11.2, 28.0)	0.7 ± 1.5 (0.1, 1.2)	0.99 (0.95, 1.00)*
belly press test	single trial	11.0 (6.2, 15.5)	0.5 ± 1.7 (-1.1, 0.1)	0.96 (0.85, 0.99)*	9.7 (3.2, 15.7)	0.4 ± 2.9 (-1.4, 0.6)	0.92 (0.74, 0.98)
	three trials	11.9 (5.7, 14.9)	0.3 ± 1.6 (-0.8, 0.3)	0.96 (0.87, 0.99)	9.4 (3.4, 14.5)	0.2 ± 1.7 (-0.8, 0.3)	0.98 (0.91, 0.99)

Abbreviations. Mean diff, mean difference between trials; LOA, limits of agreement; CI, confidence interval; ICC, intraclass correlation coefficient; abd, abduction

*$p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$

Discussion

Accurate measures of ROM and muscle strength are required when this clinical information is used for diagnostic classification purposes, to avoid undue mis-classification of patients into incorrect diagnostic categories. Further standardisation of ROM and resisted tests was undertaken in an attempt to improve measurement precision and reliability for ROM and resisted tests prior to their use in the diagnostic accuracy study. Active ROM elevation, passive ROM abduction, passive ROM external rotation (0° abduction) and resisted tests demonstrated sufficient intraexaminer reliability using more stringent test standardisation procedures to justify the use of a single trial in the larger diagnostic study. Passive ROM external and internal rotation (90° abduction) and passive ROM internal rotation demonstrated acceptable intraexaminer reliability based upon mean values from three trials.

Compared with previous intraexaminer reliability results for ROM tests (LOA 6° to 24°) (Cadogan, Laslett, et al., 2011a), the latest results demonstrate a considerable reduction in measurement variability (LOA 4.7^o to 10.8^o) for active and passive ROM tests despite similar ICC values. For measures of passive ROM external and internal rotation, results for multiple trials demonstrated reduced measurement variability (mean differences and LOA) and narrower confidence limits compared with results of single trials. These results suggest that multiple trials of passive ROM internal and external rotation may be required to achieve more consistent measures.

In the previous study, data from the peak trial during resisted tests was compared with the mean of three trials (Cadogan, Laslett, et al., 2011a), however in the current study, the first trial, rather than the peak trial was used. In order to identify the peak trial, several trials would need to be performed for all participants. Due to number of tests being considered in the diagnostic study, as well as the potential for provocation of symptoms due to the number of tests, the difference between the first trial and mean of three trials was assessed to determine whether a significant difference existed which would reduce the number of potentially provocative procedures required. The previous LOA for resisted tests (1.1 to 7.0kg) were higher than current results (0.7 to 2.3kg) showing improved repeatability and increased measurement precision with the use of more standardised test procedures and external stabilisation of the HHD when using the first trial compared with multiple trials. This suggests that even though the first trial may not have been the 'peak' trial in all cases, the difference between trials was sufficiently small that it would be unlikely to unduly affect diagnostic classification.

Passive ROM internal rotation, resisted internal rotation and force measures during the belly press test were not previously investigated, however, current results indicate sufficient measurement precision to justify the use of these measurements according to the procedures described.

This reliability study included only 12 participants, however with the exception of shorter duration of symptoms, participants demonstrated similar demographic features to participants in the previous study. Intraclass correlation coefficients are known to be adversely affected when the sample size is small, or in the presence of a narrow range of values (Shrout & Fleiss, 1979). Both factors were inherent in this small reliability study hence ICCs may not accurately reflect reliability in this sample. In the context of the diagnostic study, where absolute ROM values may be used as criteria for diagnostic classification irrespective of sample rank consistency, the absolute measurement variability was considered a more appropriate statistic for interpretation of these results. Results indicate that measurement variability was reduced to acceptable levels for this purpose following refinement of test procedures. Due to time and resource constraints it was not possible to test interexaminer reliability using the new standardised procedures, however this will be undertaken at a later date.

Conclusion

The standardised test procedures improved measurement precision for a single examiner compared with previous results. Although the sample was small, results suggest a single trial of all tests may be sufficient for the majority of tests, however multiple trials may be required for measures of passive ROM external and internal rotation performed at 90° abduction.

CHAPTER FOUR

OVERVIEW OF DIAGNOSTIC STUDY METHODS

Preface

This chapter presents the diagnostic study methodology. Analysis contained within this chapter also addresses Specific Aim 3 of the thesis:

To evaluate the prevalence of imaged pathology and assess the relationship between imaging findings and anaesthetic responses to SAB, ACJ and GHJ diagnostic blocks.

The diagnostic study methods are described in the following manuscript that was published in BMC Musculoskeletal Disorders (2011). Aspects of study methodology that were considered important but could not be included in the manuscript due to word count limitations are provided in appendices and reference is made to relevant appendices within the text.

Additional descriptive results are also presented in the manuscript that address Aim 3 of the thesis. These results relate to the prevalence of imaged pathology on standard x-ray and diagnostic ultrasound scans (all participants) and magnetic resonance arthrogram (MRA) findings in a sub-group of participants who underwent this procedure. The prevalence of positive anaesthetic responses ($\geq$80% post-injection pain relief) following the SAB, ACJ and GHJ diagnostic blocks is also reported.

Diagnostic imaging findings are frequently used to aid the clinical diagnosis of shoulder pain, however the literature review revealed a lack of information regarding the prevalence (pre-test or prior probability) of imaged pathology in a primary care population. This information is required in order to estimate the post-test probability of a specific condition being present when other clinical data are added. To date, the prevalence of imaging findings in a symptomatic and consecutive cohort of patients recruited from primary care has not been reported and the relevance of imaging findings to symptoms of shoulder pain remains unclear. This manuscript also provides analysis of the relationship between the imaged pathology and the response to diagnostic blocks.

4.1. A PROSPECTIVE STUDY OF SHOULDER PAIN IN PRIMARY CARE: PREVALENCE OF IMAGED PATHOLOGY AND RESPONSE TO GUIDED DIAGNOSTIC BLOCKS

Abstract

Background: The prevalence of imaged pathology in primary care has received little attention and the relevance of identified pathology to symptoms remains unclear. This paper reports the prevalence of imaged pathology and the association between pathology and response to diagnostic blocks into the subacromial bursa (SAB), acromioclavicular joint (ACJ) and glenohumeral joint (GHJ).

Methods: Consecutive patients with shoulder pain recruited from primary care underwent standardised x-ray, diagnostic ultrasound scan and diagnostic injections of local anaesthetic into the SAB and ACJ. Participants who reported less than 80% reduction in pain following either of these injections were referred for a magnetic resonance arthrogram (MRA) and GHJ diagnostic block. Differences in proportions of positive and negative imaging findings in the anaesthetic response groups were assessed using Fisher's test and odds ratios were calculated a for positive anaesthetic response (PAR) to diagnostic blocks.

Results: In the 208 participants recruited, the rotator cuff and SAB displayed the highest prevalence of pathology on both ultrasound (50% and 31% respectively) and MRA (65% and 76% respectively). The prevalence of PAR following SAB injection was 34% and ACJ injection 14%. Of the 59% reporting a negative anaesthetic response (NAR) for both of these injections, 16% demonstrated a PAR to GHJ injection. A full thickness tear of supraspinatus on ultrasound was associated with PAR to SAB injection (OR 5.02; $p<0.05$). Ultrasound evidence of a biceps tendon sheath effusion (OR 8.0; $p<0.01$) and an intact rotator cuff (OR 1.3; $p<0.05$) were associated with PAR to GHJ injection. No imaging findings were strongly associated with PAR to ACJ injection ($p\leq0.05$).

Conclusions: Rotator cuff and SAB pathology were the most common findings on ultrasound and MRA. Evidence of a full thickness supraspinatus tear was associated with symptoms arising from the subacromial region, while a biceps tendon sheath effusion and an intact rotator cuff were associated with an intra-articular GHJ pain source. When combined with clinical information, these results may help guide diagnostic decision making in primary care.

Introduction

Shoulder pain is a common and disabling complaint. The reported annual incidence of shoulder pain in primary care is 14.7 per 1000 patients per year (D. A. W. M van der Windt et al., 1995) with a lifetime prevalence of up to 70% (Luime, Koes, et al., 2004). Recovery from shoulder pain can be slow and recurrence rates are high with 25% of those affected by shoulder pain reporting previous episodes, and 40 to 50% reporting persisting pain or recurrence at 12-month follow-up (Croft et al., 1996; Urwin et al., 1998; D. A. W. M van der Windt et al., 1996).

The most common causes of shoulder pain in primary care are reported to be rotator cuff disorders, acromioclavicular joint (ACJ) disease and glenohumeral joint (GHJ) disorders (Mitchell, Adebajo, Hay, & Carr, 2005), with classification of these disorders based primarily upon results of clinical tests (Bot et al., 2005; Chakravarty & Webley, 1990; Chard et al., 1990; Feleus et al., 2008; Ostor et al., 2005; D. A. W. M van der Windt et al., 1995). However, inconsistent diagnostic terminology (Schellingerhout, Verhagen, Thomas, & Koes, 2008), lack of universally accepted diagnostic classification criteria (Boocock et al., 2009; Buchbinder, 1996) and poor specificity of many physical examination tests (Hegedus et al., 2008) hamper confidence in classification systems that use clinical test criteria alone.

Diagnostic imaging investigations, including shoulder x-ray and diagnostic ultrasound imaging, are increasingly being utilised by primary care practitioners to aid diagnosis (Awerbuch, 2008). More advanced imaging investigations such as magnetic resonance arthrogram (MRA) are also available, providing improved visualisation of pathologies such as glenoid labral lesions and tendon pathology (Shahabpour et al., 2008). While previous studies report the prevalence of imaging findings in the general population (Milgrom et al., 1995), specific athletic populations (Connor, Banks, Tyson, Coumas, & Alessandro, 2003; Miniaci et al., 2002), samples of convenience (Reilly, Macleod, Macfarlane, Windley, & Emery, 2006; Shubin Stein et al., 2001) or case-

control comparisons for specific shoulder pathology (Shubin Stein et al., 2006), the prevalence of imaged pathology in a prospective cohort of primary care patients suffering a current episode of shoulder pain has not been previously reported. Diagnostic decisions rely upon knowledge of prevalence of a condition in specific populations in order to estimate the likelihood of a positive 'disease' status or outcome following specific tests or investigations (Davidson, 2002). Knowledge of prevalence of imaged pathology in primary care would provide prior probability for specific conditions, thus assisting diagnostic decision-making processes and assessment as to the value of expensive or invasive investigations or interventions.

The interpretation of imaging findings can be complicated by the presence of anatomic variants (De Maeseneer et al., 2000; Sammarco, 2000) and the high prevalence of asymptomatic pathology especially in ageing populations (Milgrom et al., 1995; Shubin Stein et al., 2001). The prevalence of asymptomatic full-thickness rotator cuff tears more than doubles after the age of 50 years (Milgrom et al., 1995), and asymptomatic ACJ arthritis has been identified by magnetic resonance imaging (MRI) in 93% of individuals over the age of 30 years (Shubin Stein et al., 2001). Despite widespread use of imaging investigations in primary care, the relationship between imaging findings and symptoms has received limited attention. Diagnostic injections of local anaesthetic provide a method for determining whether symptoms arise from a specific structure (Cyriax, 1982; Neer, 1983). Following injection of local anaesthetic into an anatomical structure, any subsequent reduction in pain intensity can be measured to assess the likelihood of its involvement in the patient's symptoms (Dreyfuss et al., 1996; Strobel et al., 2003; Walton et al., 2004).

The aims of this paper were to report the prevalence of imaged shoulder pathology, and to evaluate the association between imaged pathology and a positive response to diagnostic blocks in a consecutive sample of patients with shoulder pain recruited from a primary care setting.

Methods

Study Design and Setting

The results presented in this paper formed part of a wider prospective, blinded diagnostic accuracy study in which clinical examination and imaging variables (index tests) were compared with results of diagnostic injections of local anaesthetic (reference standard) into the SAB, ACJ and GHJ. Participants were recruited consecutively from a

community-based medical centre and nine physiotherapy practices across Christchurch, New Zealand.

Ethical Approval

The New Zealand Ministry of Health Regional Ethics Committee (Upper South A) granted ethical approval in May 2008.

Participants

Consecutive patients presenting to their primary care practitioner (general practitioner (GP) or physiotherapist) for the first time with a new episode of shoulder pain (Figure 4.1), who were over 18 years of age and able to follow verbal instructions were eligible for inclusion in the study. Exclusion criteria were known fractures or dislocations around the shoulder complex, referred pain from the cervical spine, sensory or motor deficit involving the upper limb, previous surgery to the shoulder or cervical spine or contraindications to imaging or injection procedures.
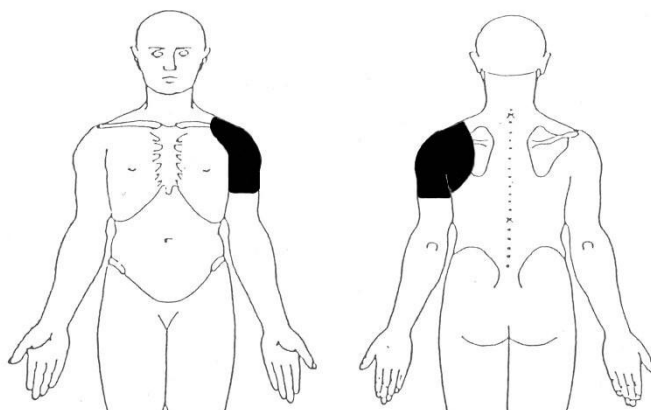
Figure 4.1. Location of primary pain required for inclusion in the study.

Procedures

A summary of all diagnostic study procedures is presented in Figure 4.2. Participants underwent a clinical examination (Appendix 7, p273) followed by a standard shoulder x-ray series, diagnostic ultrasound scan and imaging guided diagnostic injections into the SAB and ACJ. Participants reporting less than 80% reduction in pain intensity from either of these two injections were reviewed by a sports medicine physician prior to receiving an injection of local anaesthetic into the GHJ, performed as part of a contrast-enhanced MRA procedure.
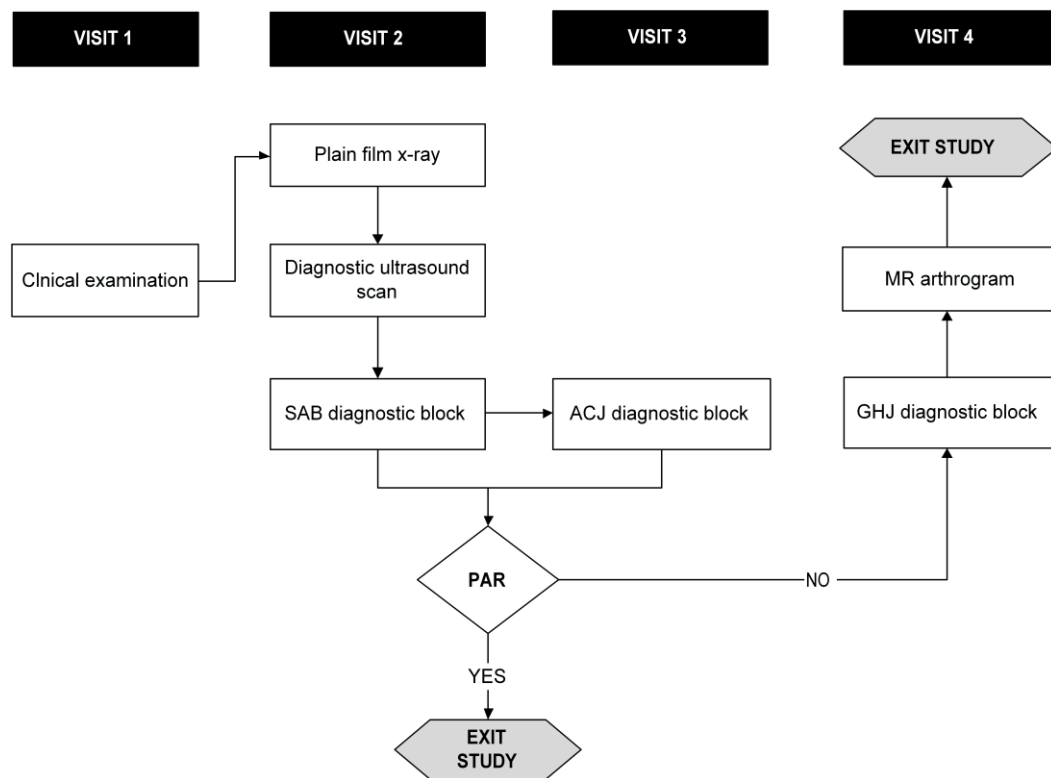
Figure 4.2. Flow chart of study procedures. SAB, subacromial bursa; ACJ, acromioclavicular joint; GHJ, glenohumeral joint, PAR, positive anaesthetic response (≥80% reduction in pain intensity); MR arthrogram, magnetic resonance arthrogram.

Treatment between procedures.

During the ethical application process it was commented that treatment could not be withheld if this was likely to result in deterioration in the patients' condition, however all participants and referring practitioners were requested to refrain from treatment for the shoulder pain wherever possible until completion of the study procedures.

X-ray and Diagnostic Ultrasound Scan

Participants underwent a standardised series of shoulder radiographs (x-ray) consisting of anterior-posterior (AP) views in neutral, external and internal rotation, axial view and outlet view (Anderson, Read, & Steinweg, 1998). X-rays were reported by experienced musculoskeletal radiologists. A standardised report form was used and radiologists recorded specific abnormalities of the ACJ, acromion, GHJ and calcific deposits. Imaging diagnostic criteria are presented in Table 4.1.

Table 4.1. *Imaging Diagnostic Criteria*

Pathology	Imaging Diagnostic Criteria
X-Ray	
Acromioclavicular joint	
arthropathy/degenerative change	joint space narrowing, subchondral sclerosis, subchondral cystic change or marginal osteophytes.
osteolysis	bony resorbtion or increased lucency in distal clavicle or acromion.
Glenohumeral joint	
arthropathy/degenerative change	joint space narrowing, subchondral sclerosis, subchondral cystic change or marginal osteophytes.
other	loose bodies, joint calcifications.
Calcification of rotator cuff components	
supraspinatus	calcific deposits adjacent to the greater tuberosity on AP-external rotation x-ray view.
infraspinatus	calcific deposits adjacent to the greater tuberosity on AP-internal rotation x-ray view.
subscapularis	calcific deposits in the anterior shoulder region on axial x-ray view.
Ultrasound[a]	
ACJ pathology	Capsular hypertrophy, cortical irregularity or osteophytes, capsular bulge, joint space narrowing or widening.
Glenohumeral joint effusion	more than 2mm between posterior glenoid labrum and posterior capsule.
Rotator cuff	
normal	normal contour, normal echogenicity.
calcification	focal increase in echogenicity with or without shadowing.
tendinosis	tendon thickening or decreased echogenicity.
intrasubstance tear	hypoechoic change not extending to articular or bursal surface.
partial thickness tear	SSp and ISp: hypoechoic change extending to either the articular or bursal surface. Subscapularis: partial fibre discontinuity.
full thickness tear	SSp and ISp: hypoechoic region extends from bursal to articular surface. Subscapularis: complete fibre discontinuity.
Subacromial bursa	
bursitis	hypoechoic fluid or effusion present and >2mm thick.
bursal thickening	≥2mm measured from deep margin of deltoid to superficial margin of supraspinatus.
"bunching"	Fluid distension of the SAB or 'buckling' of the rotator cuff during abduction
MR arthrogram[a]	
Acromioclavicular joint	
arthropathy/degenerative changes	capsular hypertrophy with or without joint space narrowing, subchondral cystic change, bone marrow oedema or osteophytes
osteolysis	bony resorption or bone marrow oedema in the distal clavicle
Rotator cuff	
normal	normal contour, normal signal
tendinosis	tendon thickening or mild increase in T2 signal
intrasubstance tear	linear increase in T2 signal which does not extend to the articular or bursal surface.
partial thickness tear	linear increase in T2 signal extending to the (bursal or articular) margins.
full thickness tear	fluid signal intensity or contrast extending from the bursal to the articular side lesion of the rotator cuff. Contrast seen in the SAB.

Subacromial bursitis	increased T2 signal within the SAB
Glenohumeral joint	
rotator interval pathology	thickening, signal change or tear involving the biceps pulley, superior glenohumeral or coracohumeral ligament, or synovitis in the rotator interval.
arthropathy/degenerative change	chondral loss, subchondral sclerosis, cystic changes, bone marrow oedema or osteophytes
labral tear	contrast extending into- or undermining the glenoid labrum, not conforming to normal variant anatomy.

Abbreviations. AP, antero-posterior view; ACJ, acromioclavicular joint; SSp, supraspinatus; ISp, infraspinatus; SAB, subacromial bursa

[a]definitions based upon accepted diagnostic criteria (McNally, 2004; Stoller, Wolf, Li, Nottage, & Tirman, 2007)

Diagnostic ultrasound scans were performed by trained and experienced musculoskeletal sonographers and reported by fellowship trained musculoskeletal radiologists. Examinations were performed using a Philips IU22 machine with a 5–12MHz linear array probe using a standardised scan procedure (Backhaus et al., 2001; McNally, 2004): 1) patient sitting with palm face up on their knee (long head of biceps tendon); 2) elbow tucked into their side with external rotation of the shoulder (subscapularis); 3) arm resting on lap in neutral rotation with elbow behind body (supraspinatus); 4) hand in the small of the back with palm facing outwards to visualise (supraspinatus); 5) hand placed on the opposite shoulder (infraspinatus, ACJ, posterior labrum and glenohumeral joint). Scanning was conducted along the line of each tendon and at 90 degrees to the tendon.

The SAB was observed during dynamic abduction and 'bunching' under the acromion and the coracoacromial ligament (CAL) was recorded. Subacromial bursal dimensions were measured from the deep margin of deltoid muscle to superficial margin of supraspinatus tendon in all cases where this distance was measurable (dimensions exceeding 1mm).

Diagnostic Injections

Subacromial bursa injection.

Participants were positioned supine with the arm in external rotation. Under aseptic conditions, a 22-gauge needle was used to inject 5mL of 1% lidocaine hydrochloride (xylocaine[TM]) into the SAB under ultrasound guidance using an anterior approach. When needle placement inside the SAB was confirmed by ultrasound, the contents of the syringe were emptied into the bursa. The radiologist recorded whether the SAB was successfully infiltrated. A video of this procedure may be viewed in the

electronic article in Additional file 3_SAB injection, compatible with Windows® Media Player software.

Acromioclavicular joint injection.

One week after the SAB injection, local anaesthetic was injected into the ACJ under fluoroscopic guidance using contrast enhancement. Participants were positioned supine with the arm in external rotation. Under aseptic conditions, a 22-gauge needle was inserted into the ACJ using a direct anterior approach. Iodinated contrast (0.5ml of Omnipaque 300 GE Healthcare) was introduced and fluoroscopic images used to confirm needle placement within the ACJ. Approximately 2mL of 1% lidocaine hydrochloride (xylocaine^{TM}) was then injected into the joint. The radiologist recorded whether the ACJ was successfully infiltrated and whether the injectate was contained within the joint. A video of this procedure may be viewed in the electronic article in Additional file 4_ACJ injection.

Glenohumeral joint injection.

Approximately one week after the ACJ injection, participants reporting less than 80% relief from both the SAB and ACJ injections underwent a GHJ arthrogram and intra-articular injection of local anaesthetic and gadolinium prior to magnetic resonance imaging (MRI). Participants were positioned supine and the GHJ injection carried out under fluoroscopic guidance as described for the ACJ injection (above) using 5mL of iodinated contrast. A mixture of 0.5mL gadolinium (0.5 mmol/ml Gd-DOTA Guerbet France) and 10mL 1% lidocaine hydrochloride (xylocaine^{TM}) was injected into the joint. The radiologist recorded whether the injectate was contained within the joint. A video of this procedure may be viewed in the electronic article Additional file 5_GHJ injection.

Determination of post-injection change in pain intensity.

Immediately prior to each injection, all participants were examined using up to six clinical tests identified as being provocative of the participant's typical symptoms during the initial clinical examination. Pre-injection pain intensity was recorded for each clinical test on a 100mm visual analogue scale (VAS) where 0mm indicated "no pain" and 100mm represented "worst imaginable pain". Tests were repeated between 5 and 15 minutes following each injection and post-injection pain intensity VAS scores recorded again. The percentage change in pain intensity (anaesthetic response) was calculated for each test [(post-injection VAS – pre-injection VAS/pre-injection VAS)*100]. The average percent change from all tests was then calculated. A post-injection reduction in

pain intensity of 80% or more was used as the criterion for a positive anaesthetic response (PAR). Participants who did not reach an average of 80% pain relief following the SAB and ACJ injection were evaluated by a sports medicine physician and referred for the MRA investigation.

Magnetic Resonance Arthrogram Imaging

Magnetic resonance imaging was obtained within 30 minutes of the GHJ injection. Imaging was performed with 3.0 Tesla General Electric-Milwaukee (GE) Signa HDxt platform running version 15 software. A conventional MR arthrography protocol was followed (Stoller et al., 2007). The patient was positioned supine with the affected arm extended alongside their body and externally rotated. Total scan time was 30 minutes including patient positioning.

Series 1: Axial oblique (Obl) T1 (Time to Echo (TE)/Time to Repeat (TR) Min full/640, Echo Train Length (ETL) 4, receive bandwidth (BW) 41.67, slice thickness (ST)/slice gap (SG) 3mm/0mm, field of view (FOV) 16cm, frequency/phase matrix (Freq/Phase) 320/320, number of excitations (NEX) 2).

Series 2: Coronal Obl T1 with Fat saturation (FS) (TE/TR min full/480, ETL 4, BW 35.71, ST/SG 3.5mm/0mm, FOV 16cm, Freq/Phase 288/288, NEX 3).

Series 3: Coronal Obl Proton Density (PD) FS Forced Recovery Fast Spin Echo (FRFSE XL), TE/TR 40/3660, ETL 10, BW 31.25, ST/SG 3.5mm/0mm, FOV 16cm, Freq/Phase 320/320, NEX 2).

Series 4: Sagittal T2 Obl FRFSE XL (TE/TR 65/3060, ETL 16, BW 41.67, ST/SG 3mm/0mm, FOV 14cm, Freq/Phase 320/320, NEX 2). The GE 8-channel high definition shoulder coil was used for all sequences.

For the abduction-external rotation (ABER) sequence the arm was raised so that the palmar aspect of the hand was resting under the patient's occiput and their elbow as close to the table surface as possible. Series 5: Coronal T1 FS Obl ABER (TE/TR min full/480, ETL 4, BW 31.25, ST/SG 3.5mm/0mm, FOV 18cm, Freq/Phase 256/256, NEX 2). The GE 6 channel flex phased array coil was used for the ABER sequence.

Blinding

The investigator performing the clinical examination and pre- and post-injection clinical tests (AC) was blinded to all diagnostic and treatment information from referring practitioners and to the results of imaging procedures. Sonographers and

radiologists were blinded to all clinical information prior to the x-ray, ultrasound scans and MRA procedure, and were blinded to results of anaesthetic response to injections.

Sample Size Considerations

Sample size was estimated using methods described by Flahault et al., (2005) (Flahault, Cadilhac, & Thomas, 2005). Sample size was calculated for the diagnostic sub-group with the lowest expected prevalence (ACJ). The minimal acceptable lower confidence limit was set at 0.75 and expected sensitivity/specificity were both set at 0.90. A review of sample size estimates after the first 100 cases indicated lower than expected prevalence of PAR to ACJ diagnostic block and sample size was adjusted in order to maintain precision of diagnostic estimates.

Statistical Analysis

The prevalence of imaged pathology and response to each of the diagnostic blocks are reported as frequency and percentages. Contingency tables (2x2) were constructed and Fisher's exact test was used to compare proportions of positive and negative imaging findings in the anaesthetic response groups for each diagnostic injection procedure. *P*-values of ≤ 0.05 were used to indicate statistical significance. Odds ratios (OR) and 95% confidence intervals (CI) for PAR to diagnostic blocks were calculated. Statistical Package for the Social Sciences (SPSS) version 17.0 (IBM® Corporation 2010) was used for the analysis.

Due to the known limitations of VAS scales for measuring change in pain intensity when pre-injection pain levels are low (<20mm) (Bogduk, 2004b), only cases where pre-injection pain intensity exceeded 20mm were included in the analysis of anaesthetic response to diagnostic injections. Average percentage change in pain intensity was calculated for the index tests with positive integers indicating increased post-injection pain intensity, and negative integers indicating decreased post-injection pain intensity.

Results

Participants

A total of 208 participants were included in the study between July 2009 and June 2010. Details of progression of participants through the study and dropout explanations are presented in Figure 4.3. Demographic information for those included in the study is

presented in Table 4.2. There were no significant differences between those included and those excluded from the study with respect to age or gender. Symptom duration was shorter (median 2 weeks; IQ range 4 weeks) in participants excluded from the study (Mann-Whitney $p<0.001$). There were no significant differences in demographic characteristics between the total sample and the sub-group who received the GHJ injection as part of the MRA procedure ($p>0.05$).
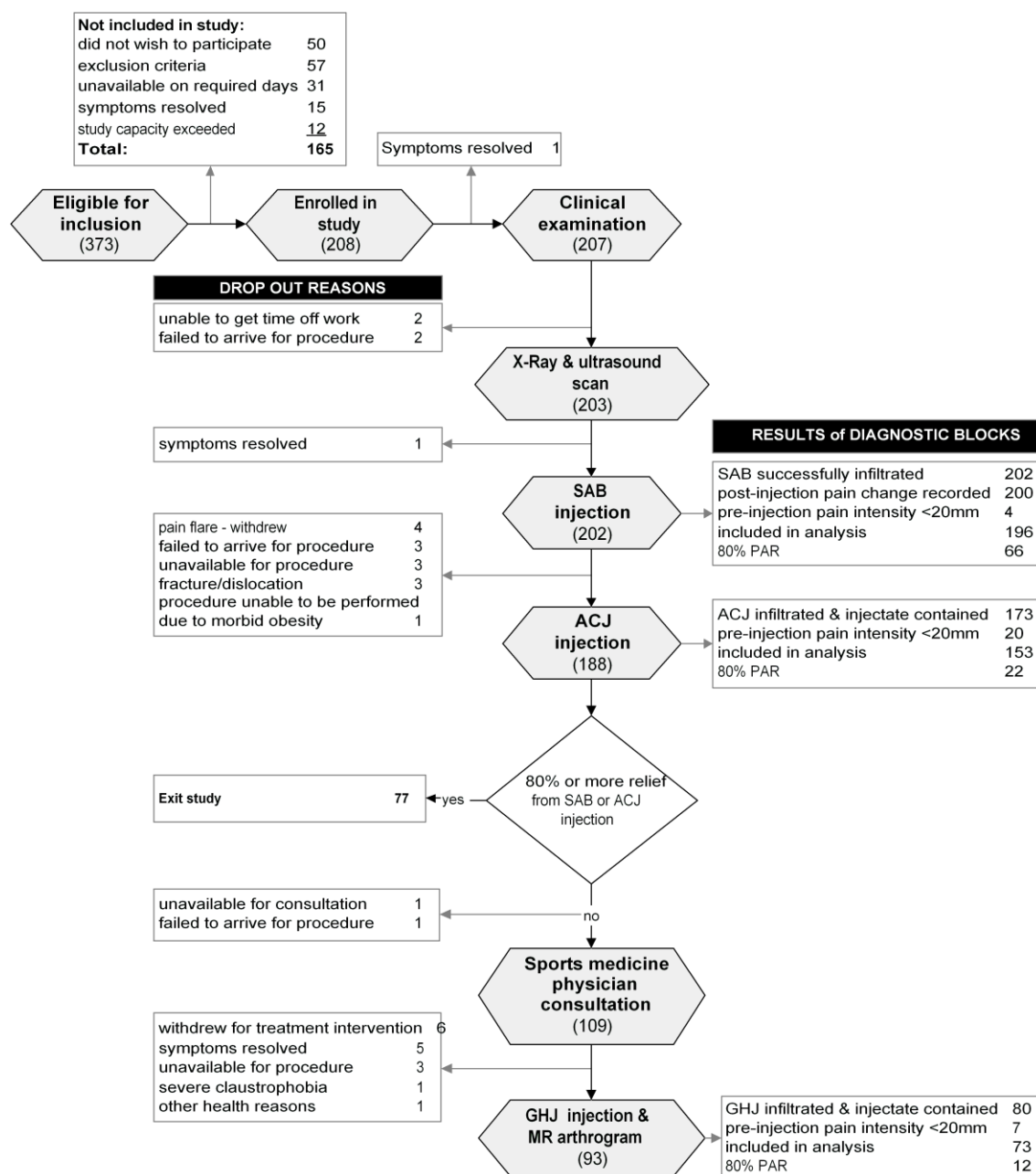
Figure 4.3. Flow chart showing results of diagnostic accuracy study. Diagram summarises reasons for exclusion of participants, dropout explanations, results of diagnostic block procedures and adverse reactions. SAB, subacromial bursa; PAR, positive anaesthetic response (≥80% pain relief); ACJ, acromioclavicular joint; GHJ, glenohumeral joint; MR arthrogram, magnetic resonance arthrogram; IV, intravenous.

Table 4.2. *Participant Demographics*

Characteristics	All participants (n=208)		MRA group (n=93)	
	Mean (SD)	Range	Mean (SD)	Range
Age (years)	42 (14)	18-81	42 (14)	18-81
Height (cm)	172 (10)	147-199	172 (10)	151-198
Weight (kg)	80.6 (18.0)	50.3-189.0	82.3 (15.8)	52.7-125.3
Symptom duration (weeks)[a]	7 (13)	0-175	7 (13)	0-175
Worst pain previous 48 hours (100mm VAS)	62 (23)	3-100	63 (24)	3-100
Average pain previous 48 hours (100mm VAS)	37 (22)	1-100	37 (24)	1-100
	n (%)		n (%)	
Male gender	107 (51)		53 (57)	
Right hand dominant	110 (53)		79 (85)	
Dominant arm affected	110 (53)		48 (52)	
ACC Claim	193 (93)		86 (93)	
Referrals				
physiotherapist	203 (98)		89 (96)	
general practitioner	5 (2)		4 (4)	
Employment status				
in paid employment	166 (80)		76 (82)	
on modified duties due to shoulder pain	18 (9)		10 (11)	
off work due to shoulder pain	7 (3)		4 (4)	
not currently employed	41 (20)		17 (18)	
Co-existent medical conditions	70 (34)		33 (36)	
Current smoker	39 (19)		18 (20)	

Abbreviations. MRA, magnetic resonance arthrogram; SD, standard deviation; VAS, visual analogue scale: ACC, Accident Compensation Corporation.
[a]symptom duration was not normally distributed. Figures presented are median (IQ range).

Prevalence of Imaged Pathology

X-ray and ultrasound scan

The prevalence of the pathologies identified on x-ray and ultrasound are presented in Figures 4.4 and 4.5. Acromioclavicular joint (Figure 4.6a) and GHJ pathology were the most common x-ray findings (both 17%) and calcification involving the rotator cuff was reported in 13% of participants (Figure 4.6b).

Rotator cuff pathology was the most prevalent pathology on ultrasound (50%), with supraspinatus the most commonly affected rotator cuff component, accounting for 86 of the 102 cases (85%) of rotator cuff pathology. Tears were the most common pathology affecting supraspinatus accounting for 52% of all supraspinatus pathology. Intrasubstance tears were the most common type of tear accounting for 51% of all supraspinatus tears (Figure 4.7a). Calcification was the most common finding in both infraspinatus (59%) and subscapularis (69%) compared with 39% in supraspinatus.

Prevalence of SAB pathology was 31% and bursal thickening (dimensions exceeding 2mm) was reported in 23% of participants (Figure 4.7b). Bunching of the SAB under the acromion was observed in 84 participants (43%) (Figure 4.7c), and this was associated with reproduction of symptoms in 72 participants (86% of cases in which bunching was observed). Bunching under the CAL was observed in 51 of the 94 cases (54%) in which this was assessed, and was associated with reproduction of symptoms in 40 participants (78% of cases in which bunching was observed) (Figure 4.8).
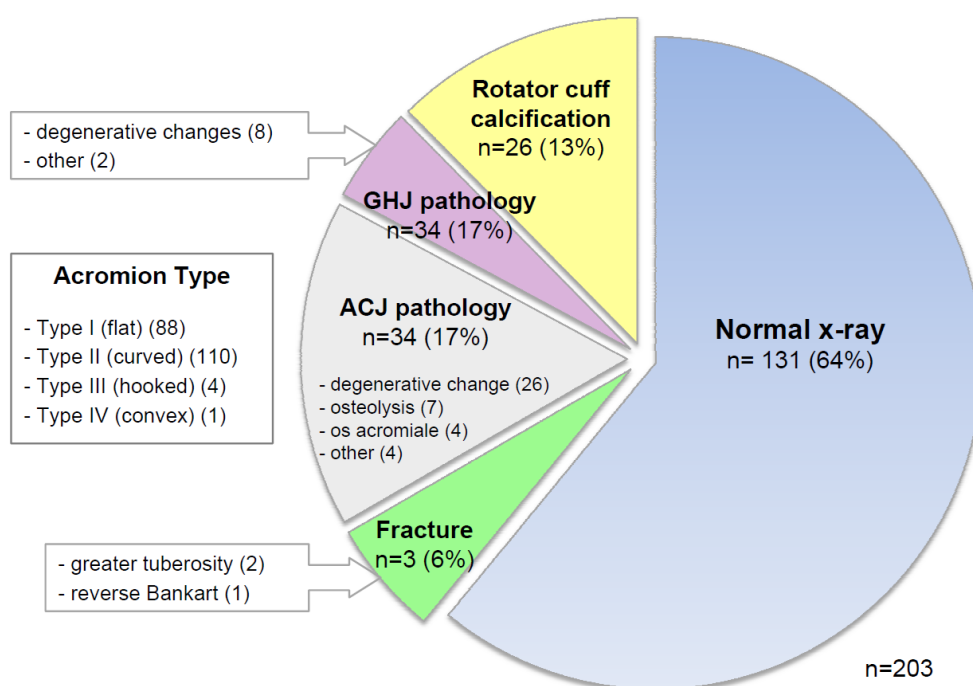
Figure 4.4. Prevalence of pathology identified on x-ray. n, number of cases; ACJ, acromioclavicular joint; GHJ, glenohumeral joint.
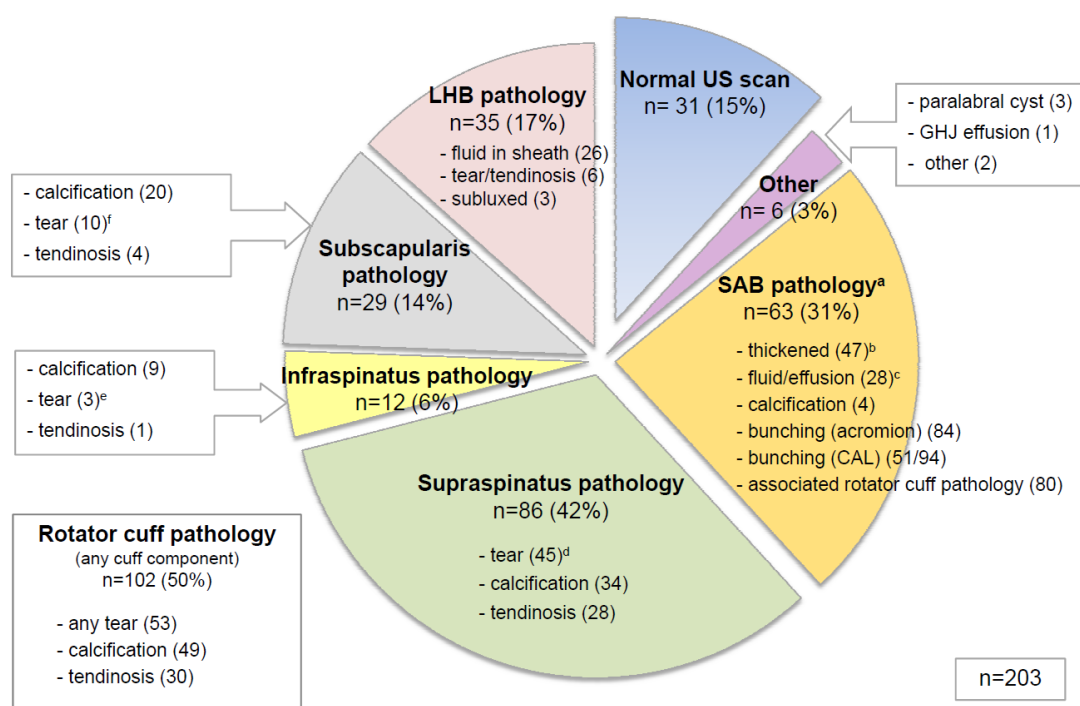
Figure 4.5. Prevalence of pathology identified on ultrasound scan.
(n), number of cases; US, ultrasound; GHJ, glenohumeral joint; SAB, subacromial bursa; CAL, coracoacromial ligament; LHB, long head of biceps tendon.
[a]Subacromial pathology: any one of three present; dimension ≥2mm, fluid/effusion or calcification; [b]Subacromial bursa dimensions: <1mm (71); 1-2mm (82); 2-3mm (42); >3mm (5); [c]Subacromial bursal effusion associated with full thickness rotator cuff tear (7). [d]Supraspinatus tears: intrasubstance (23); partial thickness-bursal surface (4); partial thickness-articular surface (8); full thickness (10). [e]Infraspinatus tears: intrasubstance (1); partial thickness (1); full thickness (1). [f]Subscapularis tears: intrasubstance (5); partial thickness (4); full thickness (1).

Figure 4.6. Shoulder x-ray images of ACJ pathology and rotator cuff calcification. Figure shows a) AP x-ray view in external rotation showing degenerative acromioclavicular joint changes (white arrow); b) outlet view showing calcification in line with the infraspinatus tendon (black arrow).
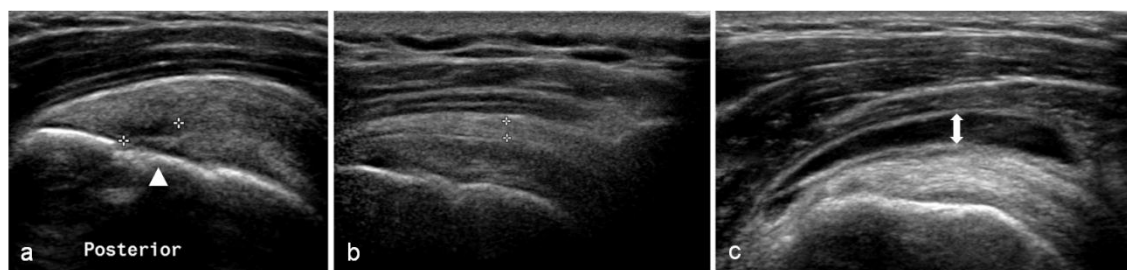
Figure 4.7. Ultrasound scan images of subacromial bursa and supraspinatus pathology. Figure shows a) hypoechoic region (between calipers) indicating an intrasubstance tear within posterior fibres of supraspinatus (longitudinal view) overlying the head of humerus (white arrowhead); b) thickened subacromial bursa (calipers); c) bunching of the SAB (white arrow) under the acromion during dynamic abduction.

	Acromion (n=195)	Coracoacromial ligament (n=94)
■ symptomatic bunching	37%	42%
■ asymptomatic bunching	6%	12%
■ no bunching	57%	46%
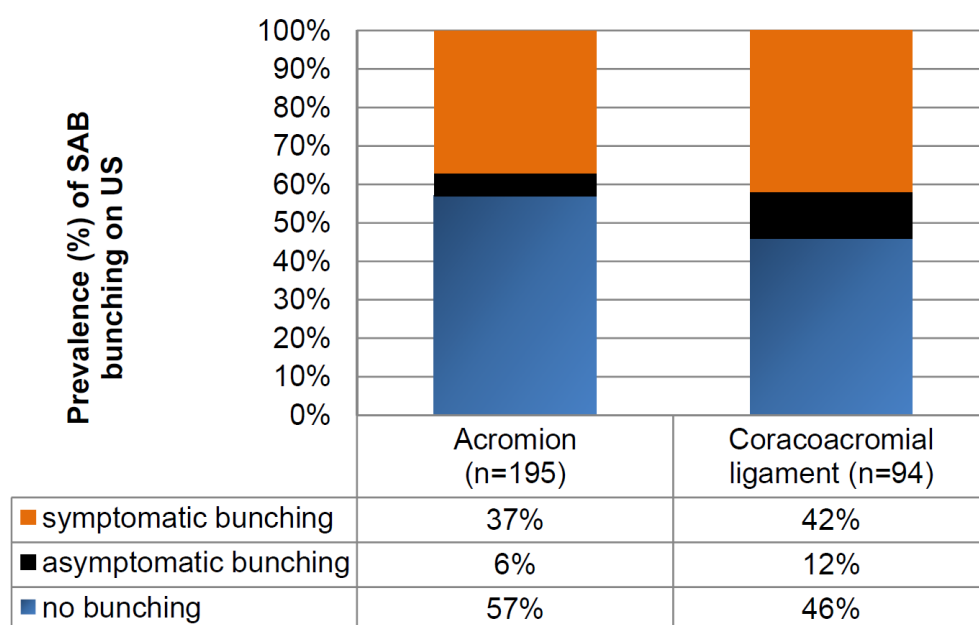
Figure 4.8. Prevalence of subacromial bursa bunching under the acromion and coracoacromial ligament on ultrasound during dynamic abduction. SAB, subacromial bursa; US, ultrasound; CAL, coracoacromial ligament. Percentages are in reference to the number of cases in which bursal bunching was assessed (acromion n=195; CAL n=94).

Magnetic resonance arthrogram.

The prevalence of MRA findings is shown in Figure 4.9. Only one case was reported as "normal" (no abnormality reported) and 74% of cases demonstrated multiple pathologies. The most commonly reported MRA finding overall was SAB pathology (76%) with subacromial bursitis reported in 68 participants (73%) (Figure 4.10a). Rotator cuff pathology affected at least one of the rotator cuff components in 65% of cases. Supraspinatus was the most frequently affected component of the rotator cuff (85% of all rotator cuff pathology) and tears were the most common pathological finding in all rotator cuff components accounting for 41 of the 61 cases (67%) of rotator cuff pathology. Partial thickness tears involving the articular surface were the most

common type of supraspinatus tear identified (34% of all supraspinatus tears) (Figure 4.10b). GHJ pathology (63%) and ACJ pathology (59%) were also highly prevalent with rotator interval pathology (GHJ) and degenerative ACJ changes (Figure 4.10c) (both 55%) the most common findings. Glenoid labrum tears were present in 47% of all participants who received the MRA and were associated with paralabral cysts in 10 cases (23%). Suprascapular nerve compression was associated with paralabral cysts in two cases (2%) (Figure 4.10d).

Prevalence of Anaesthetic Response to Diagnostic Blocks

The anaesthetic response profiles for the diagnostic injections are presented in Appendix 8 (p275). There were no observable differences in the frequency of imaged pathology between those in whom post-injection pain intensity increased compared with cases in which a post-injection decrease in pain was reported for any of the diagnostic block procedures. Infiltration of the SAB was confirmed in all cases and a PAR (≥80% pain relief) was reported by 66 participants (34%) following the SAB injection.

Average ACJ injection volume was 2.1mL (SD 0.7mL) and 22 of the 153 participants (14%) in whom the injectate was contained within the ACJ and whose pre-injection pain intensity exceeded 20mm on the 100mm VAS scale reported an 80% PAR. Ninety three participants received the GHJ injection as part of the MR arthrogram procedure and an 80% PAR was reported by 12 of the 75 participants (16%) in whom the injectate was contained within the GHJ and pre-injection pain intensity exceeded 20mm.

Association Between Imaged Pathology and Response to Diagnostic Blocks

Imaging variables associated with PAR to diagnostic block ($p \le 0.05$) and demonstrating a magnitude of association OR greater than 2.0 are summarised in Table 4.3. Results for all other x-ray and ultrasound variables are presented in Appendix 9 (p276) (SAB and ACJ injection), and Appendix 10 (p279) (GHJ injection). Results for all other MRA variables are presented in Appendix 11 (p281).
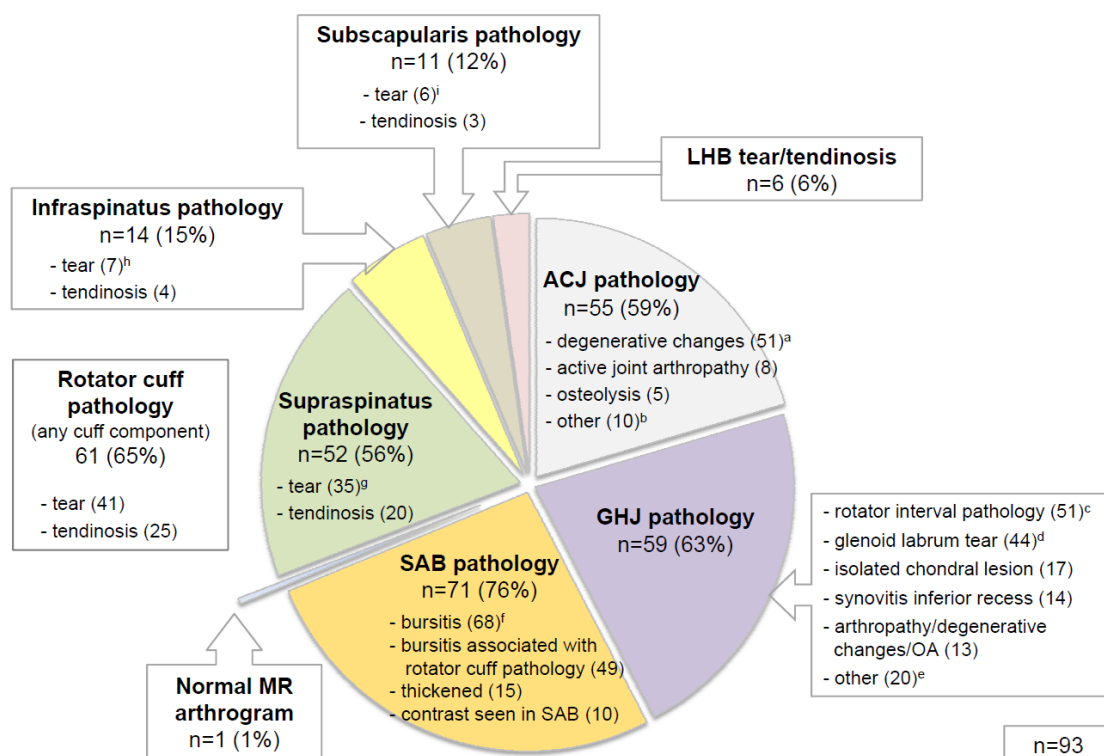
Figure 4.9. Prevalence of pathology identified on MR arthrogram. *Abbreviations:* (n), number of cases; LHB, long head of biceps tendon; ACJ, acromioclavicular joint; GHJ, glenohumeral joint; OA, osteoarthritis; SAB, subacromial bursa.

[a]ACJ degenerative changes: mild (28), moderate (18), severe (5).

[b]Acromioclavicular joint pathology – other: os acromiale (2), unfused acromial ossification centre (1), acromial spur (4), widened joint space/subluxation (2), synovitis (1).

[c]Rotator interval pathology: coracohumeral or superior glenohumeral ligament thickening (40), rotator interval synovitis (39), biceps pulley, coracohumeral or superior glenohumeral ligament tear (13).

[d]Glenoid labrum tear: isolated labral tear (5), associated pathology present (39), SLAP tear (20), SLAP Type II (17), Type III (2), Type IV (1), anterior-inferior tear (9), semi- or full circumferential tear (7), posterior-superior tear (1), other tear (9), paralabral cyst (10), paralabral cyst causing suprascapular nerve compression (2).

[e]Glenohumeral joint pathology – other: bony irregularity humeral head without marrow oedema (12), Hill-Sachs lesion (3), intra-articular/osseous body (3), ganglion cyst between coracoacromial and coracohumeral ligaments (1), greater tuberosity fracture (1).

[f]Subacromial bursitis: mild (52), moderate (12), severe (4).

[g]Supraspinatus tears: intrasubstance (11), partial thickness-bursal surface (5), partial thickness articular surface (12), full thickness (7).

[h]Infraspinatus tears: intrasubstance (4), partial thickness (3), full thickness (0).

[i]Subscapularis tears: intrasubstance (4), partial thickness (0), full thickness (2).

Figure 4.10. MR arthrogram images of shoulder pathology. Figures show a) subacromial bursitis – coronal PD fat saturated image showing region of hyperintensity in the subacromial bursa (black arrow); b) partial thickness, articular surface supraspinatus tear (white arrow) - coronal T1 fat saturated image showing contrast extending into the supraspinatus tendon; c) ACJ degenerative changes (white arrow) -coronal PD fat saturated image; d) type III SLAP tear (white arrow) with contrast filling a paralabral cyst (black arrow) which extended into the supraglenoid and suprascapular notch causing neural compression -coronal PD fat saturated image.

A full thickness supraspinatus tear identified by ultrasound imaging was associated with PAR to SAB injection (OR 5.0, $p \leq 0.05$). None of the imaging variables were strongly associated with PAR to ACJ injection ($p > 0.05$). The strongest association of any imaging variable with diagnostic block was the association between biceps tendon sheath effusion identified on ultrasound and PAR to GHJ injection (OR 8.0; $p < 0.01$). A tear of the rotator cuff reported on ultrasound was negatively associated with a PAR to GHJ injection ($p < 0.05$). When recoded, an 'intact' rotator cuff on ultrasound demonstrated an OR of 1.3 for a PAR.

Table 4.3. *Imaging Variables Associated with Positive Anaesthetic Response to Diagnostic Blocks*

Pathology identified on imaging	Pathology identified (total cases) (n)	% with pathology present reporting PAR	% with pathology absent reporting PAR	OR (95% CI)	Fishers test (*p*-value)
SAB injection (PAR n=66)					
X-ray: type 3 acromion	4	75	33	6.2 (0.64, 61.23)	0.109
X-ray: os acromiale	4	75	33	6.1 (0.63, 60.25)	0.112
X-ray: supraspinatus calcification	16	56	31	2.8 (1.00, 7.97)	0.054
US: supraspinatus calcification	33	49	31	2.1 (1.00, 4.55)	0.068
US: supraspinatus FTT	10	70**	32	5.0 (1.25, 20.11)	0.033
ACJ injection (PAR n=22)					
X-ray: ACJ pathology	21	14	16	2.1 (0.69, 6.52)	0.189
US: supraspinatus tear PTT (articular surface)	8	0	17	2.1 (0.39, 11.05)	0.323
US: LHB tendinosis	3	0	16	3.1 (0.27, 35.39)	0.374
GHJ injection (PAR n=12)					
US: no rotator cuff tear	19	21**	0	1.3 (1.11, 1.46)	0.029
US: supraspinatus tendinosis	11	27	14	2.3 (0.51, 10.30)	0.374
US: subscapularis tendinosis	3	33	15	2.8 (0.23, 33.27)	0.421
US: biceps tendon sheath effusion	13	46**	10	8.0 (2.02, 31.72)	0.004
MRA: ACJ pathology	46	20	11	2.0 (0.50, 8.23)	0.516
MRA: osteolysis lateral clavicle	5	40	15	3.9 (0.58, 26.58)	0.187
MRA: contrast seen in SAB	6	33	15	2.9 (0.47, 17.99)	0.254

Abbreviations. PAR, positive anaesthetic response (≥80% post-injection pain intensity reduction); OR, odds ratio; CI, confidence interval; SAB, subacromial bursa; US, ultrasound; FTT, full thickness tear; ACJ, acromioclavicular joint; PTT, partial thickness tear; LHB, long head of biceps; GHJ, glenohumeral joint; MRA, magnetic resonance arthrogram
Note. Variables are presented that were associated with a PAR ($p \leq 0.05$) or with an odds ratio of ≥2.0
Percentages do not total 100% as these represent proportion of participants with or without pathology on imaging (row percentages in contingency table) in the PAR group. Negative anaesthetic response group results (column percentages) are not presented
**$p \leq 0.05$.

Discussion

This is the first report of the prevalence of imaged pathology and anaesthetic responses to diagnostic injection into the SAB, ACJ and GHJ in a sample of primary care patients with shoulder pain. Estimates of the likelihood of symptomatic pathology being present that affect these sites will increase or decrease as details from the history and physical examination are added to the imaging findings, but prior probability (prevalence) of these conditions in the population of interest is the necessary baseline and starting point (Davidson, 2002). This study provides the prior probability data for specific pathologies and pain sources at the 80% pain reduction level in a sample of primary care patients. This knowledge may help inform clinical decisions regarding treatment interventions, the use of advanced imaging or specialist referral.

Prevalence of Imaged Pathology

X-ray and diagnostic ultrasound scan.

Shoulder x-rays were reported as 'normal' in 64% of cases however the detection of three unsuspected fractures in the study population highlights the value of x-ray as a screening tool. The prevalence of calcification identified on x-ray (13%) was similar to previous reports (10%) (Speed & Hazleman, 1999).

Subacromial bursa pathology was a common ultrasound finding (31%) in this symptomatic sample. The bursal dimension of ≥2mm, calcification or bursal fluid or effusion was used to classify 'SAB pathology'. Opinions vary regarding the dimension (thickness) at which the normally thin hypoechoic line of the SAB is regarded as pathological. Some have suggested the ability to view and measure the SAB at all represents pathological thickening (Farin et al., 1990), others consider more than 2mm thickness to be pathological (Kolla & Motamedi, 2007; Naredo et al., 2002; M. Van Holsbeeck & Strouse, 1993) and some suggest SAB thickness compared with the unaffected side irrespective of bursal dimension to be of more clinical relevance (Tsai et al., 2007). Recent theories question whether SAB thickening is even pathological, proposing it may be the result of adaptation to repeated overhead activity (Awerbuch, 2008). Variable agreement (kappa 0.50 to 0.89) has also been reported between musculoskeletal ultrasound experts for identification of SAB pathology on ultrasound (Bruyn et al., 2009; Bruyn et al., 2010; Le Corroller et al., 2008; Naredo et al., 2006) with most disagreements relating to variations in dynamic assessment and judgement of SAB fluid as being normal or pathological (Naredo et al., 2006). Technicalities surrounding the ultrasound diagnosis of SAB pathology, lack of expert consensus upon the dimension at which the SAB is considered pathological and the poor understanding of the relationship between SAB histopathology and imaging findings mean that the reported prevalence of SAB pathology on ultrasound is likely to vary. Bursal bunching was also identified in a high proportion of participants, however bunching was asymptomatic in 14% (acromion) and 22% (CAL) of cases in which bunching was observed. This highlights the need to correlate imaging findings with clinical symptoms when considering the diagnosis of 'subacromial impingement'.

Magnetic resonance arthrogram.

Magnetic resonance arthrogram findings in the subgroup of participants receiving this investigation, revealed a high prevalence of multiple pathologies (74%), similar to previous reports (77%) in an asymptomatic primary care population (Sher et al., 1995).

In the participants who received the MRA, SAB and ACJ pathology were reported respectively in 76% and 59% of participants, all of whom had previously been classified as 'non-responders' at the 80% pain relief level following injection of local anaesthetic into these structures. Marrow oedema on MRI has been reported as a reliable indicator of symptomatic ACJ pathology (Shubin Stein et al., 2006). Our study identified eight cases (9%) of active ACJ arthropathy with marrow oedema in participants who had previously demonstrated a NAR to ACJ injection, however the inability of the local anaesthetic to penetrate to the level of subchondral bone, thereby classifying those participants as 'non-responders' to ACJ injection, represents a likely explanation for this result.

Rotator cuff pathology was reported in more than half of participants on both ultrasound and MRA with rotator cuff tears identified in 26% and 44% of participants with the respective imaging procedures. Although no primary care imaging studies are available for direct comparison, these results are similar to previous reports of the prevalence of rotator cuff tears in asymptomatic populations on ultrasound (Milgrom et al., 1995) and MRI (Sher et al., 1995). Of interest was the higher number of intrasubstance tears involving infraspinatus, and partial thickness (articular surface) supraspinatus tears identified on MRA compared with the number identified on ultrasound imaging, despite the smaller sample number in this subgroup. While identification of an intrasubstance tear on MRA is unlikely to alter management at primary care level unless it is associated with more serious pathology, partial thickness tears of the rotator cuff are reported to be of prognostic significance due to the high proportion that increase in size or progress to full thickness tears if left untreated (Yamanaka & Matsumoto, 1994). Ultrasound imaging has previously demonstrated only moderate pooled sensitivity (72%) for detection of partial thickness rotator cuff tears compared with MRI or surgery (Ottenheijm et al., 2010). Variable agreement among experts on the presence of partial thickness rotator cuff tears on ultrasound (kappa 0.63; 88% to 92% agreement) has also been reported (Le Corroller et al., 2008; Middleton, Teefey, & Yamaguchi, 2004; Naredo et al., 2006). Results of MRI scans have been shown to alter clinical decisions regarding management of rotator cuff tears in the orthopaedic setting (Sher et al., 1998) and MRA may therefore be indicated at the primary care level if there is clinical suspicion of rotator cuff disruption in the presence of equivocal ultrasound findings.

The prevalence of intra-articular GHJ pathology on MRA in this sub-group of participants was also high (63%) with rotator interval pathology (55%) and glenoid labral tears (47%) the most common findings. However, despite the high prevalence of GHJ pathology in this study, only 16% of individuals were classified as responders to the GHJ injection at the 80% pain relief level. During the MRA procedure, contrast was introduced into the GHJ through the region of the rotator interval and in some participants the appearance of contrast in this region on subsequent MRI films may have been difficult to distinguish from mild rotator interval pathology. Glenoid labral tears are frequently associated with other extra-articular pathology such as rotator cuff tears (Alasaarela, Takalo, Tervonen, Hakala, & Suramo, 1997; Bussières et al., 2008; D. Chang, Mohana-Borges, Borso, & Chung, 2008; G. Walch, Boileau, Noel, & Donell, 1992), and the rotator interval also has complex pathoanatomic relationships with supraspinatus, subscapularis and the long head of biceps tendon (Pradhan & Itoi, 2001). The high proportion of multiple pathology and low GHJ PAR rate in this study may be partially explained by the concurrent involvement of extra-articular structures.

Association between Imaging Findings and Anaesthetic Response

Participants with full thickness tears of supraspinatus identified by ultrasound imaging were more likely to experience a PAR to SAB injection than those without a full thickness tear. Full thickness supraspinatus tears affect the SAB-rotator cuff interface and infiltration of the torn cuff with anaesthetic through this disruption is the likely explanation for this finding. The small proportion of PAR among those with an intrasubstance supraspinatus tear (intact margins) reported on ultrasound supports this theory, however none of the four cases in which bursal-surface supraspinatus tears were identified were classified as responders to the SAB injection. None of the imaging variables were strongly associated with PAR to ACJ injection. The high prevalence of asymptomatic degenerative changes, particularly in individuals older than 30 years, (93%) (Shubin Stein et al., 2001) may explain this result.

A long head of biceps tendon sheath effusion on ultrasound was significantly related to a PAR to GHJ injection. The biceps tendon sheath is a synovial extension of the GHJ capsule and may therefore be indicative of a GHJ effusion resulting from intra-articular GHJ pathology or systemic inflammatory disease. A biceps tendon sheath effusion on ultrasound has been shown to be more sensitive than arthrography for detection of intra-articular GHJ pathology (Middleton, Reinus, & Totty, 1985). It is also a common finding in those suffering rheumatoid arthritis (Iagnocco, Coari, Leone, &

Valesini, 2003; Iagnocco et al., 2006) and has been found to be predictive of degenerative GHJ arthritis and polymyalgia rheumatica (Alasaarela et al., 1997; Lange et al., 2000). In the current primary care study, half the participants with a biceps tendon sheath effusion reported on ultrasound were classified as positive 'responders' to the GHJ diagnostic block at the 80% pain reduction standard. The likely explanation for the PAR is the anaesthetisation of synovial tissue within the GHJ. Although this finding may implicate an intra-articular pain source, it is a non-specific result and further imaging investigations such as MRI or laboratory tests would be required to identify the specific pathology responsible for the synovial effusion. The magnitude of association of the biceps tendon sheath effusion on ultrasound with PAR to GHJ injection seen in this study (OR 8.00), and a lower 95% confidence limit of 2.0 suggest this finding may be of value in the primary care setting when considering further imaging investigation, laboratory testing or referral for higher levels of care.

Participants with an intact rotator cuff on ultrasound also demonstrated a higher proportion of PAR to GHJ injection ($p<0.05$) than those in whom a rotator cuff tear was identified. This could imply that in participants with a rotator cuff tear, the tear itself may have been more symptomatic than any co-existent intra-articular GHJ pathology resulting in the NAR to GHJ diagnostic block. Although the OR for PAR to GHJ injection in the presence of an intact rotator cuff on ultrasound was small (1.27), the CI did not include 1.0, and could represent a clinically meaningful increase in the likelihood of a PAR since the prevalence of this imaging finding was high (74%) (Peat & Barton, 2005). Current guidelines advocate ultrasound imaging only when a major rotator cuff tear is suspected when surgery may be considered as a treatment option (Arrol et al., 2004, p. 3). However, these results may provide additional justification for the use of diagnostic ultrasound imaging in the primary care setting to inform decisions regarding further investigations for intra-articular GHJ pathology in the presence of an intact rotator cuff and relevant clinical findings.

Limitations of the Study

The definition of 'accident' in the context of participant 'claim status' in this study is influenced by New Zealand's' unique Accident Compensation Corporation legislation. Although the majority of participants included in this study had a current ACC claim, this does not necessarily imply a significant degree of trauma, and complaints included many less severe conditions with low levels of functional disability. Those whose shoulder pain is not covered by an ACC claim may, however,

be less likely to present for medical assessment and may be under-represented in this study. Due to the cost of the MRA procedures it was not possible for every participant to undergo this procedure, and several participants with high and low levels of pain intensity withdrew from the study prior to the MRA representing a potential source of selection bias in this subgroup of participants.

Conclusion

Rotator cuff and SAB pathology were the most common findings on both ultrasound and on MRA in this primary care cohort. A full thickness supraspinatus tear seen on ultrasound was associated with subacromial pain according to the 80% pain relief criterion, while ultrasound findings of a biceps tendon sheath effusion and an intact rotator cuff were associated with pain arising from the GHJ in a subgroup of participants. Results provide the prior probability of imaged pathology, and when combined with clinical examination findings may inform decisions in primary care regarding treatment interventions and the need for advanced diagnostic imaging or specialist referral.

CHAPTER FIVE

PREDICTORS OF A POSITIVE RESPOSE TO SUBACROMIAL BURSA DIAGNOSTIC BLOCK

Preface

This chapter is the first of three chapters in which diagnostic accuracy results are reported.

This chapter relates to Specific Aims 2 and 4 of the thesis:

To estimate the diagnostic accuracy of clinical examination findings for identifying a predominant subacromial source of shoulder pain defined by a positive response to diagnostic block.

To evaluate the added diagnostic value of imaging findings for predicting a positive response to SAB diagnostic block.

This chapter (Chapter 5) consists of two manuscripts reporting the diagnostic accuracy of clinical examination and imaging findings for predicting a positive response to subacromial bursa diagnostic block. The first manuscript presents clinical examination features that were associated with a positive anaesthetic response (PAR) following the SAB diagnostic block and reports the diagnostic accuracy of these clinical features for predicting a PAR. This manuscript has been submitted to the Journal of Rehabilitation Medicine. In the second manuscript, imaging findings that were associated with an 80% SAB PAR were added to the clinical examination features. Diagnostic accuracy was then re-calculated, and the added value of these imaging findings for ruling-in and ruling-out a PAR was assessed.

To reflect clinical practice where patients typically undergo a clinical examination prior to referral for diagnostic imaging a regression model was derived in which the strongest clinical examination predictors of a PAR were identified (model 1) and diagnostic accuracy of these variables was calculated. Model 1 results are presented in the first manuscript. The strongest imaging predictors for a PAR were then identified and added to the clinical examination prediction model, to derive the best combination of both clinical and imaging features for predicting a PAR (model 2), and diagnostic accuracy re-calculated. Results for model 2 are presented in the second manuscript.

Both sections in this chapter are presented as individual manuscripts hence they contain some repetition of methodology.

5.1. CLINICAL PREDICTORS OF A POSITIVE RESPONSE TO GUIDED DIAGNOSTIC BLOCK INTO THE SUBACROMIAL BURSA

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Taylor, S. Clinical predictors of a positive response to guided diagnostic block into the subacromial bursa (*submitted*).

Abstract

Objective: To identify the strongest clinical predictors of positive response to injection of local anaesthetic into the subacromial bursa (SAB).

Design: Prospective, cohort, diagnostic validity design.

Participants: Consecutive patients with shoulder pain recruited from primary care physiotherapy and general medical practices.

Methods: All participants underwent a standardised clinical examination (index test) followed by a diagnostic injection of xylocaine^{TM} into the SAB (reference standard test) performed under ultrasound guidance. Clinical examination variables associated with a positive anaesthetic response (PAR) ($\geq$80% post-injection reduction in pain intensity) were identified ($p<0.20$) and diagnostic accuracy was calculated.

Results: A PAR was reported by 34% of participants. Strain injury (adjusted odds ratio (AOR) 2.3), anterior shoulder pain (AOR 2.3) and absence of pain with external rotation at 90° abduction (AOR 3.9) were the strongest clinical predictors of PAR (100% specificity when all three were positive). Combinations of nine clinical variables demonstrated 100% sensitivity (95% CI 0.95, 1.00) for a PAR when at least one of the findings was not present, and 97% specificity (95% CI 0.92, 0.99) for a PAR when six or more findings were present.

Conclusion: Combinations of these clinical tests may assist the clinician to differentiate subacromial pain from other shoulder conditions and guide selection of targeted pain management interventions.

Introduction

Shoulder pain is a common and disabling complaint with a reported prevalence in the general population of at least 16% (Urwin et al., 1998), and up to 34% in those over the age of 65 years (Chakravarty & Webley, 1990). Shoulder pain is also frequently associated with medical conditions including diabetes mellitus (Garcilazo, Cavallasca, & Musuruana, 2010), and is reported by up to 91% of those with rheumatoid disease (Olofsson, Book, & Jacobsson, 2003; Petersson, 1986). Subacromial disorders including subacromial bursa (SAB) pathology, rotator cuff disease and rotator cuff tears are the most commonly reported shoulder disorders, accounting for up to 70% of shoulder pain seen in primary care practice (D. A. W. M van der Windt et al., 1995). It is generally accepted that the SAB is the main source of pain in rotator cuff disease due to its anatomic location and rich nociceptive innervation (Gotoh, Hamada, Yamakawa, & al, 2001; Ide, Shirai, & Ito, 1996; Vangsness, Ennis, Taylor, & Atkinson, 1995).

The SAB maybe affected by a number of conditions including primary synovitis (bursitis) (Farin et al., 1990), crystal deposition, calcific loose bodies (Salzman, Lillegard, & Butcher, 1997), rotator cuff disease or may occur secondary to repeated mechanical 'impingement' against the acromial arch (Neer, 1983). Specific pain management interventions are also advocated in the management of painful bursal conditions including corticosteroid injections, surgical bursectomy for inflammatory bursal pathology (Blaine et al., 2005; Voloshin et al., 2005), and barbotage procedures for calcific lesions (De Conti et al., 2010). Early detection of painful bursal pathology would therefore facilitate timely application of appropriate treatment to reduce the considerable functional disability and adverse health and psychosocial consequences associated with ongoing shoulder pain (Bostrom, Harms-Ringdahl, & Nordemar, 1995; Eberhardt & Fex, 1995; Ostor et al., 2005). The success of any treatment intervention however, is dependent upon identification of the SAB as the pain source in the first instance.

Subacromial disorders may be difficult to differentiate from other sources of shoulder pain due to the complex regional anatomy, and the similar clinical presentations of different shoulder disorders (Dinnes et al., 2003). The majority of previous studies assessed the diagnostic ability of isolated physical examination tests, reporting poor specificity of these tests for identifying subacromial pathology (Calis et al., 2000; Dinnes et al., 2003; Hegedus et al., 2008; Park et al., 2005), and their limited ability to differentiate between early stage "impingement" (bursal pathology) and more advanced

rotator cuff disease (Park et al., 2005). In clinical practice, diagnosis is rarely based upon the result of a single tests, and several methods of evaluating the diagnostic accuracy of combinations of clinical tests have been reported, including clinical prediction models (Litaker et al., 2000) and combinations based upon minimum numbers of positive clinical findings (at least one, two etc) (Laslett, McDonald, et al., 2006; Murrell & Walton, 2001). To our knowledge the two methods of interpreting diagnostic accuracy for combinations of clinical data have not previously been compared in the same cohort to determine which method provides the largest improvement in post-test probability of a positive 'case'.

The majority of previous diagnostic studies used surgery as the reference standard procedure, and while this provides visualisation of pathology, it does not take into account whether the observed pathology is the primary source of symptoms. Diagnostic injections of local anaesthetic into the subacromial regions are considered the reference standard test for identification of subacromial pain (Neer, 1983), with marked reduction in post-injection pain intensity following injection of local anaesthetic into the SAB being indicative of a positive anaesthetic response (PAR) and a likely subacromial pain source. In addition to providing valuable diagnostic information, a PAR may also provide an indication of the therapeutic value of targeted pain management interventions such as corticosteroid injections. Identification of clinical examination findings with the strongest predictive ability for a subacromial pain source would assist the clinician in more efficient differentiation of subacromial pain from other shoulder conditions, facilitate appropriate additional investigative pathways for subacromial pathology, and enable more timely application of appropriate treatment interventions.

The aim of this study was to identify clinical examination predictors of a positive anaesthetic response (PAR) to a guided subacromial diagnostic block into the SAB, and determine which combinations of clinical examination variables provide the highest level of diagnostic accuracy for a PAR.

Methods

This study formed part of a wider prospective, blinded diagnostic accuracy study in which clinical examination and imaging variables (index tests) were compared with results of guided diagnostic injection of local anaesthetic (reference standard) into the SAB, acromioclavicular joint (ACJ) and glenohumeral joint (GHJ). Participants were recruited from community-based medical and physiotherapy practices across

Christchurch, New Zealand. The New Zealand Ministry of Health Regional Ethics Committee (Upper South A) granted ethical approval for the study. Informed consent was gained from all participants prior to participation in the study and the rights of all participants were protected.

Participants

Consecutive patients over the age of 18 years, presenting to their general practitioner or physiotherapist for the first time with a new episode of shoulder pain and with the ability to follow verbal instructions were eligible for inclusion in the study (Figure 4.1). Exclusion criteria were known fractures or dislocations around the shoulder complex, referred pain from the cervical spine, sensory or motor deficit involving the upper limb, previous surgery to the shoulder or cervical spine, or contraindications to injection procedures.

Procedures

Clinical examination.

All included participants completed self-report questionnaires consisting of SF-8™ health survey (Appendix 12, p282) (Ware, Kosinski, Dewey, & Gandek, 2001), Shoulder Pain and Disability Index (SPADI) (Appendix 13, p284) (Roach, Budiman-Mak, Songsiridej, & Lertratanakul, 1991) and Fear Avoidance Beliefs Questionnaire (FABQ) (Appendix 14, p285) (Gordon Waddell, Newton, Henderson, Somerville, & Main, 1993). This was followed by a standardised clinical examination including medical history (Appendix 15, p287), symptom chart (Appendix 16, p288), patient history (Appendix 17, p290) and physical examination (Appendix 18, p292). The clinical examination was conducted by an experienced clinician (A.C). A full list of clinical examination variables and response criteria are presented in Appendix 7 (p273).

The physical examination consisted of the following tests: active range of motion (ROM) of the cervical spine (Maitland, 1986), inspection for swelling or muscle atrophy, recording of symptom responses associated with passive ROM (Cadogan, Laslett, et al., 2011a) and resisted muscle tests, orthopaedic tests selected according to evidence for reported diagnostic accuracy (Hegedus et al., 2008) and performed as originally described; Hawkins-Kennedy test (Hawkins & Kennedy, 1980), drop-arm test (Codman, 1934), empty can test (F. Jobe & Moynes, 1982), external rotation lag sign (Hertel et al., 1996), Speed's test (Gill, El Rassi, Bahk, Castillo, & McFarland, 2007), apprehension-relocation test (F. W. Jobe & Kvitne, 1989) and pain responses to

palpation of the shoulder region (Mattingly & Mackarey, 1996). During the physical examination, those tests provocative of typical pain were identified for use in pre- and post-injection testing. Indeterminate results of clinical examination tests were recorded and coded as missing data.

Subacromial bursa diagnostic block.

For the subacromial diagnostic block (reference standard) procedure, participants were positioned supine with the arm in external rotation. Under aseptic conditions, a 22-gauge needle was used to inject 5mL of 1% lidocaine hydrochloride (xylocaine[TM]) into the SAB under ultrasound guidance using an anterior approach. When needle placement inside the subacromial bursa was confirmed by ultrasound, the contents of the syringe were emptied into the bursa.

Immediately prior to the injection, all participants were examined using up to six tests identified during the clinical examination as being provocative of typical symptoms. Pre-injection pain intensity was recorded for each clinical test on a 100mm visual analogue scale (VAS) where 0mm indicated "no pain" and 100mm represented "worst imaginable pain". Tests were repeated between 5 and 15 minutes following the diagnostic block and post-injection pain intensity VAS scores recorded again. The average change in pain intensity from all clinical tests was then calculated. A positive anaesthetic response was determined by 80% or more post-injection reduction in pain intensity (80% PAR). This is similar to the criteria for PAR used in other studies involving diagnostic blocks (Dreyfuss et al., 1996; Laslett, McDonald, et al., 2006; Strobel et al., 2003) and represents a high level of confidence that the target structure is a major contributor to symptoms.

Blinding.

The investigator performing the clinical examination and pre- and post-injection clinical tests (AC) was blinded to any diagnostic or treatment information from referring practitioners. The radiologist who performed the SAB diagnostic block was blinded to any clinical information and to the results of pre-injection provocative clinical testing.

Sample size estimation.

Sample size was estimated using methods for estimates for diagnostic accuracy studies described by Flahault et al.(Flahault et al., 2005) with the minimal acceptable lower confidence limit set at 0.75 and expected sensitivity/specificity both set at 0.90, with adjustment following sub-group analysis of the first 100 cases to maintain precision of confidence interval estimates.

Statistical Analysis

The Fisher exact test (dichotomous variables) and univariate logistic regression analyses (continuous variables) were performed for all demographic, self-report questionnaires and clinical examination variables for PAR to SAB diagnostic block using Statistical Package for the Social Sciences (SPSS) version 17.0 (IBM® Corporation 2010). Variables demonstrating univariate association with PAR to SAB diagnostic block at the $p \leq 0.20$ level were included in multiple logistic regression analyses and stepwise backward variable elimination was performed using Akaike's Information Criterion (AIC) (Akaike, 1974) to derive the best prediction model. Multiple regression analysis was carried out using "R" statistical software (R Development Core Team, 2010). The goodness of fit for the model was assessed using the Hosmer-Lemeshow test (Hosmer & Lemeshow, 2000). Due to the known limitations of VAS scales for measuring change in pain intensity when pre-injection pain levels are low (<20mm) (Bogduk, 2004b), only cases where pre-injection pain intensity exceeded 20mm were included in the analysis.

Diagnostic accuracy statistics including sensitivity, specificity, predictive values, positive likelihood ratios (+LR) and negative likelihood ratios (-LR) and 95% confidence intervals (CI) were then calculated (Appendix 19, p297). These were used to assess the discriminatory ability of the prediction model, and for the combinations of clinical variables associated with PAR to SAB diagnostic block ($p \leq 0.20$) according to minimum number of variables present. Confidence Interval Analysis software (Bryant, 2000) was used for calculation of diagnostic accuracy statistics.

Results

A total of 373 patients were referred to the study between July 2009 and June 2010 resulting in 208 participants being included. Demographic data for those included in the study are presented in Table 5.1. Symptom duration was significantly less (Mann-Whitney $p<0.001$) in those excluded from the study (median 2 weeks; IQ range 4 weeks). The mean time between clinical examination and the SAB diagnostic block was 4 days (SD, 3 days; range 1-19). Details of progression of participants through the study, drop-out explanations and adverse events are presented in Figure 5.1.

Table 5.1. *Demographic Information*

Demographic information	All participants (N=202)		PAR Group (n=69)	NAR Group (n=133)
	Mean (SD)	Range	Mean (SD)	Mean (SD)
Age (years)	42 (14)	18-81	42 (12)	42 (15)
Height (cm)	172 (10)	147-199	171 (9)	172 (10)
Weight (kg)	80.6 (18.0)	50.3-189.0	80.2 (21)	81 (17)
Symptom duration (weeks)[*]	7 (13)[*]	0-175	7 (14)[*]	7 (12)[*]
VAS (worst)	62 (23)	3-100	62 (22)	64 (24)
VAS (average)	37 (22)	1-100	36 (18)	37 (23)
VAS (best)	9 (18)	0-98	7 (13)	10 (20)
SF8 physical component score	44 (8)	23-61	44 (8)	44 (8)
SF8 mental component score	52 (9)	27-66	53 (8)	52 (9)
SPADI pain score (%)	50 (22)	0-100	50 (21)	51 (22)
SPADI disability score (%)	30 (23)	0-96	28 (22)	31 (22)
SPADI total (%)	38 (21)	0-98	36 (20)	39 (21)
FABQ physical activity score (%)	64 (22)	0-100	62 (23)	66 (22)
FABQ work score (%)[a]	27 (23)	0-81	26 (23)	27 (24)
FABQ total score (%)[a]	41 (19)	0-87	40 (18)	41 (19)
% male gender	51		47	55
% right hand dominant	87		88	87
% dominant arm affected	53		52	53
% ACC claim	93		92	92
% physiotherapist referrals	98		99	97
Employment status				
% in paid employment	80		82	80
% on modified duties	9		9	9
% off work	3		0	5
% co-existent medical conditions	34		32	35
% smoker	19		19	19

Abbreviations. PAR, positive anaesthetic response (≥80% post-injection reduction in pain intensity); NAR, negative anaesthetic response (<80% reduction in post-injection pain intensity); VAS, 100mm visual analogue pain score in previous 48 hours; SPADI, Shoulder Pain & Disability Index; FABQ, Fear Avoidance Beliefs Questionnaire; ACC, Accident Compensation Corporation; GP, general practitioner.
[a]only cases 'in paid employment' used in analysis.
[*]variable not normally distributed; median (interquartile range) are presented

Two hundred and seven participants completed the clinical examination. Variables for which missing data exceeded five percent included 'family history of shoulder pain' (15% 'unsure'), atrophy in the supraspinous or infraspinous fossa (9% indeterminate) and painful arc abduction (1% 'unsure' if typical symptoms were reproduced; 11% had insufficient active ROM abduction). Frequency distributions of clinical examination results for the PAR and NAR groups are presented in Table 5.2.

Two hundred and two participants received the SAB diagnostic block with needle placement being confirmed within the SAB prior to injection in all 202 cases. Post-injection change in pain intensity was obtained from 200 participants. Four cases were excluded in which pre-injection pain intensity was less than 20mm on the VAS scale, resulting in 196 cases being included in the analysis. An 80% PAR was reported by 66

of the 196 (34%) cases following the SAB injection. Eleven participants (6%) reported a post-injection increase in pain intensity.
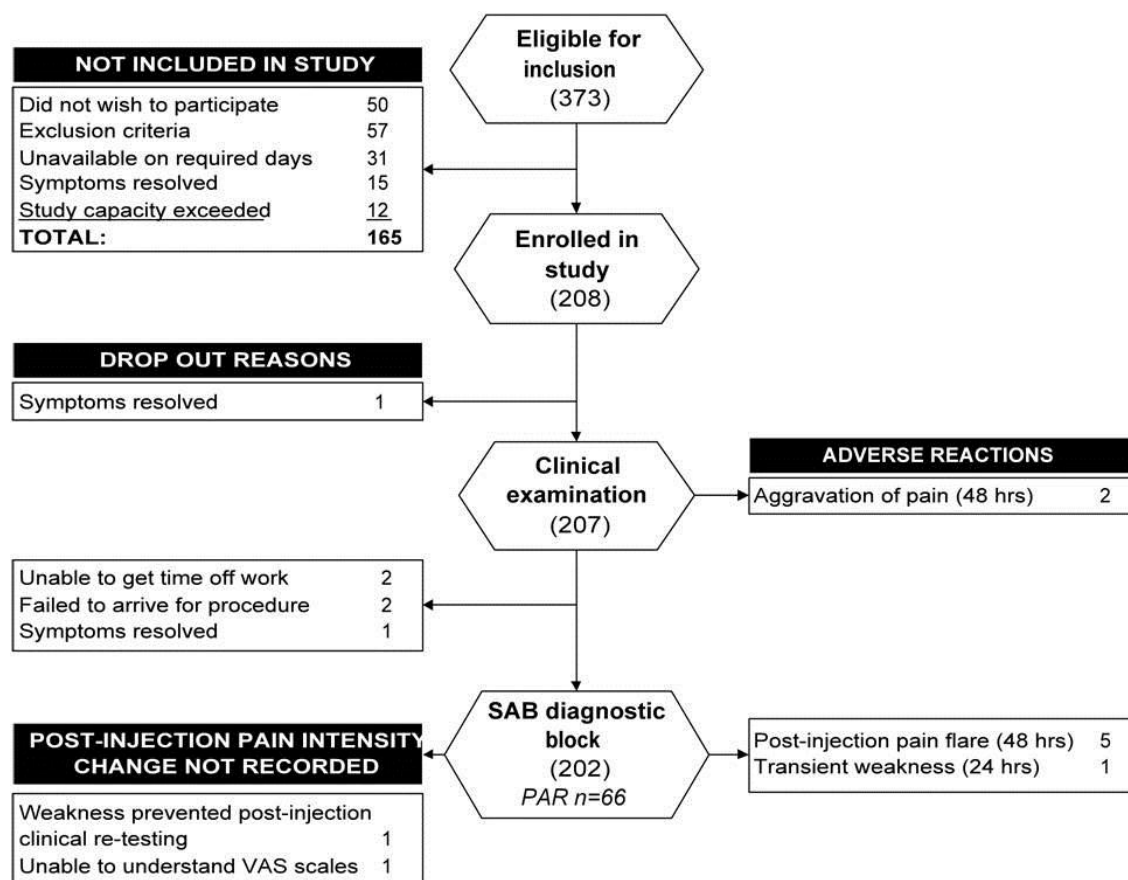
Figure 5.1. Flow chart of progression through study. This figure describes the progression of participants through the study, dropout explanations and adverse reactions. *Abbreviations:* SAB, subacromial bursa; PAR, positive anaesthetic response.

Prediction Model

No demographic or self-report variables were associated with a PAR to SAB diagnostic block ($p \leq 0.20$). Table 5.3 presents univariate odds ratios (OR), contingency cell counts and diagnostic statistics for potential clinical examination predictors associated with PAR to SAB diagnostic block ($p \leq 0.20$). The most efficient clinical examination predictors of a PAR to SAB diagnostic block were anterior shoulder pain (adjusted odds ratio (AOR) 2.3), strain mechanism of injury (AOR 2.3) and the absence of symptom provocation during passive ROM external rotation (at 90° abduction) (AOR 3.9). Hosmer-Lemeshow statistics indicated the goodness of fit of the model was adequate ($\chi^2_6 = 3.24$, $p = .778$). Diagnostic accuracy of combinations of prediction model variables is presented in Table 5.4. Highest sensitivity (0.40; 95% CI 0.29, 0.52) was observed when at least one of the three variables was not present, and highest specificity (1.00; 95% CI 0.97, 1.00) was observed when all three clinical variables were present.

Table 5.2 *Distribution of Main Clinical Examination Findings*

	Total number of positive tests (n)		
	All participants (N=196)	PAR group (n=66)	NAR group (n=130)
History			
Past history of shoulder pain	64	22	42
Family history of shoulder pain	37	13	24
Mechanism of onset			
traumatic	74	17 [*]	57
strain	81	36 [**]	45
repetitive	22	9	13
unknown	18	3	15
Pain location			
anterior	63	28 [*]	35
superior	31	10	21
lateral shoulder/arm	57	17	40
posterior	10	4	6
Pain aggravated by overhead activity	187	63	124
Referred pain extending below the elbow	28	9	19
Nature of pain constant/intermittent	61	21	40
Night pain disturbs sleep	100	37	63
Unable to sleep on the affected side	105	39	66
Physical examination			
Cervical spine pain on testing	100	36	64
AROM elevation[a] – symptoms reproduced	163	52	111
AROM HBB – symptoms produced	136	40	96
Painful arc abduction	101	35	66
Resisted tests – symptoms reproduced	172	60	112
any resisted test[b]	172	60	112
resisted abduction or external rotation	154	50	104
resisted internal rotation	93	33	60
PROM – symptoms reproduced with testing			
glenohumeral abduction	153	45 [*]	108
external rotation (0^0) abduction	136	45	91
external rotation (90^0) abduction	147	39 [***]	108
internal rotation (90^0) abduction	107	31	76
cross-body adduction (IR)	130	38 [*]	92
Orthopaedic tests			
Hawkins-Kennedy test	125	38	87
drop-arm test	20	8	12
empty can test (pain or weakness)	163	57	106
external rotation lag sign	7	3	4
Speed's test	125	36	89
apprehension/relocation (pain)	73	22	51
Palpation – typical symptoms reproduced			
greater tuberosity	105	33	72
lesser tuberosity	81	22	59
long head of biceps tendon	103	34	69

Abbreviations. PAR, positive anaesthetic response (≥80% post-injection reduction in pain intensity); NAR, negative anaesthetic response (<80% post-injection reduction in pain intensity), AROM, active range of motion; PROM, passive range of motion; IR, internal rotation; ER, external rotation; SAB, subacromial bursa; max, maximum; CAL, coracoacromial ligament.
[a]elevation through flexion
[b]symptoms reproduced with any of: resisted abduction, external rotation or internal rotation
Significant difference in frequency of positive clinical examination tests between PAR and NAR groups:*$p<0.05$; ** $p<0.01$; *** $p≤0.001$.

Table 5.3. *Diagnostic Accuracy of Individual Clinical Examination Variables for a Positive Response to Subacromial Bursa Diagnostic Block*

Dichotomous variables	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)	AOR (95% CI)
		Cell counts					Diagnostic accuracy				Odds ratios	
Strain injury	36	30	45	85	0.55 (0.43, 0.66)	0.65 (0.57, 0.73)	0.44 (0.34, 0.55)	0.74 (0.65, 0.81)	1.58 (1.13, 2.17)	0.70 (0.51, 0.91)	2.3[**] (1.2, 4.2)	2.3[*] (1.2, 4.4)
Anterior shoulder pain	28	38	35	95	0.42 (0.31, 0.54)	0.73 (0.65, 0.80)	0.44 (0.33, 0.57)	0.71 (0.63, 0.78)	1.58 (1.05, 2.33)	0.79 (0.61, 0.98)	2.0[*] (1.0, 3.7)	2.3[*] (1.2, 4.5)
Unable to sleep on affected side	39	22	66	62	0.64 (0.51, 0.75)	0.48 (0.40, 0.57)	0.37 (0.29, 0.47)	0.74 (0.64, 0.82)	1.24 (0.95, 1.58)	0.75 (0.50, 1.06)	1.7 (0.9, 3.1)	
HBB - asymptomatic	25	40	28	96	0.39 (0.28, 0.51)	0.77 (0.69, 0.84)	0.47 (0.34, 0.60)	0.71 (0.62, 0.78)	1.70 (1.08, 2.65)	0.80 (0.63, 0.97)	2.1[*] (1.1, 4.1)	
PROM GHJ abd asymptomatic	20	45	19	108	0.31 (0.21, 0.43)	0.85 (0.78, 0.90)	0.51 (0.36, 0.66)	0.71 (0.63, 0.77)	2.06 (1.19, 3.54)	0.81 (0.67, 0.96)	2.5[*] (1.2, 5.2)	
PROM ER90^o asymptomatic	26	39	20	108	0.40 (0.29, 0.52)	0.84 (0.77, 0.90)	0.57 (0.42 ,0.70)	0.74 (0.66, 0.80)	2.56 (1.56, 4.21)	0.71 (0.56, 0.86)	3.6[***] (1.8, 7.2)	3.9[***] (1.9, 8.0)
PROM IR90^o asymptomatic	32	31	47	76	0.51 (0.39, 0.63)	0.62 (0.53, 0.70)	0.41 (0.30, 0.52)	0.71 (0.62, 0.79)	1.33 (0.95, 1.84)	0.80 (0.59, 1.04)	1.7 (0.9, 3.1)	
PROM CB adduction (IR) asymptomatic	27	38	33	92	0.42 (0.30, 0.54)	0.74 (0.65, 0.81)	0.45 (0.33, 0.58)	0.71 (0.62, 0.78)	1.57 (1.04, 2.36)	0.79 (0.62, 0.98)	2.0[*] (1.1, 3.7)	
Negative Hawkins-Kennedy test	27	38	36	87	0.42 (0.30, 0.54)	0.71 (0.62, 0.78)	0.43 (0.31, 0.55)	0.70 (0.61, 0.77)	1.42 (0.95, 2.10)	0.83 (0.64, 1.03)	1.7 (0.9, 3.2)	

Abbreviations. TP, true positives; FN, false negative; FP, false positive; TN, true negatives; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; *p*, Fishers test *p*-value; OR, univariate odds ratio; AOR; multivariate adjusted odds ratio; HBB, hand-behind-back; PROM, passive range of motion; GHJ abd, glenohumeral joint abduction; ER90^o, external rotation (at 90° abduction); IR90^o, internal rotation (at 90° abduction); CB, cross body; IR, internal rotation

Note. variables were selected based upon association with an 80% positive anaesthetic response (PAR) (*p*≤0.20)

Total cell counts are less than 196 for some variables due to missing data.

Significant association between clinical examination variable and PAR to SAB diagnostic block:

*$p<0.05$; ** $p<0.01$; *** $p≤0.001$.

Table 5.4. *Diagnostic Accuracy of Prediction Model Variables for a Positive Response to Subacromial Bursa Diagnostic Block*

Number of positive clinical findings	Contingency Table Counts				Diagnostic Accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)
One of three	26	39	62	66	0.40 (0.29, 0.52)	0.52 (0.43, 0.60)	0.30 (0.21, 0.40)	0.63 (0.53, 0.72)	0.83 (0.57, 1.15)	1.16 (0.89, 1.50)	0.7 (0.4, 1.2)
Two of three	22	43	18	110	0.34 (0.24, 0.46)	0.86 (0.79, 0.91)	0.55 (0.40, 0.69)	0.72 (0.64, 0.78)	2.41 (1.40, 4.13)	0.77 (0.62, 0.91)	3.3 (1.6, 6.7)
Three of three	6	59	0	128	0.09 (0.04, 0.19)	1.00 (0.97, 1.00)	1.00 (0.61, 1.00)	0.68 (0.62, 0.75)	~ (1.45, 444.00)*	0.91 (0.84, 0.98)*	3.2 (2.6, 3.9)

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; ~, infinity.

Note: Clinical examination tests: strain mechanism of injury, anterior shoulder pain and passive range of motion external rotation (at 90° abduction) does not reproduce typical symptoms.

Total cell counts are less than 196 for some variables due to missing data.

*0.5 added to cells to estimate confidence intervals

Diagnostic Accuracy

Diagnostic accuracy of combinations of all clinical examination variables that were associated with a PAR to SAB diagnostic block (*p*<0.20) are presented in Table 5.5. Sensitivity was highest (1.00; 95% CI 0.95, 1.00) and -LR lowest (0.00, 95% CI 0.00, 1.79) when at least one clinical finding was not present. Specificity (1.00; 95% CI 0.97, 1.00) and +LR (infinity; 95% CI estimates 1.71, 509.00) were highest when at least seven clinical findings were present. Area under the ROC (0.686; 95% CI 0.598, 0.774) indicated the optimal diagnostic point was represented by four positive clinical findings (sensitivity 0.55, specificity 0.70).

Discussion

The ability to accurately identify those patients likely to report a PAR to subacromial diagnostic block can inform diagnostic decision making regarding referral for further investigation or specialist consultation and guide the selection of targeted pain management interventions such as corticosteroid injection (Cummins, Sasso, & Nicholson, 2009). Accurate identification of subacromial pain may also guide treatment selection within conservative management programmes targeted at commonly reported causes of subacromial pain including scapula dyskinesis (Burkhart, Morgan, & Kibler, 2003; B. W. Kibler, 1998) and humeral head stability (Warner, Micheli, Arslanian, Kennedy, & Kennedy, 1990). The consequences of delayed diagnosis of subacromial pain include prolonged diagnostic processes with extended periods of pain and declining functional ability and a delay in implementation of appropriate management with resulting adverse effects on treatment outcome.

Diagnostic Accuracy of the Prediction Model

The clinical prediction model identified three variables that were able to rule-in an 80% PAR to SAB diagnostic block with 100% specificity (95% CI 0.97, 1.00) when all three variables were positive (strain mechanism of injury, pain primarily located in the anterior shoulder region and when typical shoulder symptoms were not provoked during passive ROM external rotation performed at 90° abduction). However, only 3% of the primary care participants fitted this criterion, limiting the number of patients to which this model could be applied when generalized to this clinical setting. Despite the low prevalence, when present, these three findings could provide justification for the use of more invasive or expensive investigation or treatment interventions, and may of more diagnostic value in specialist settings where the prevalence of painful bursal pathology is likely to be higher.

Table 5.5. *Diagnostic Accuracy of Combinations of Clinical Examination Variables for a Positive Response to Subacromial Bursa Diagnostic Block*

Number of positive clinical examination findings	Contingency cell counts				Diagnostic accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95%CI)
One or more	66	0	123	4	1.00 (0.95, 1.00)	0.03 (0.01, 0.08)	0.35 (0.29, 0.42)	1.00 (0.51, 1.00)	1.03 (1.02, 1.09)	0.00 (0.00, 1.79)	0.7 (0.6, 0.7)
Two or more	61	3	92	26	0.95 (0.87, 0.98)	0.22 (0.16, 0.30)	0.40 (0.33, 0.48)	0.90 (0.74, 0.96)	1.22 (1.09, 1.38)	0.21 (0.07, 0.62)	5.8 (1.7, 19.8)
Three or more	50	15	62	61	0.77 (0.66, 0.86)	0.50 (0.41, 0.58)	0.45 (0.36, 0.54)	0.80 (0.70, 0.88)	1.53 (1.22, 1.91)	0.47 (0.28, 0.73)	3.3 (1.7, 6.6)
Four or more	36	27	34	92	0.57 (0.45, 0.69)	0.73 (0.65, 0.80)	0.51 (0.40, 0.63)	0.77 (0.69, 0.84)	2.12 (1.48, 3.03)	0.59 (0.42, 0.78)	3.6 (1.9, 6.8)
Five or more	23	40	14	112	0.37 (0.26, 0.49)	0.89 (0.82, 0.93)	0.62 (0.46, 0.76)	0.74 (0.66, 0.80)	3.29 (1.83, 5.90)	0.71 (0.57, 0.85)	4.6 (2.2, 9.8)
Six or more	13	51	4	125	0.20 (0.12, 0.32)	0.97 (0.92, 0.99)	0.77 (0.53, 0.90)	0.71 (0.64, 0.77)	6.55 (2.34, 18.48)	0.82 (0.70, 0.91)	8.0 (2.5, 25.6)
Seven or more	7	58	0	129	0.11 (0.05, 0.21)	1.00 (0.97, 1.00)	1.00 (0.65, 1.00)	0.69 (0.62, 0.75)	~ (1.71, 509)*	0.89 (0.82, 0.97)[*]	3.2 (2.6, 4.0)
Eight or more	3	62	0	130	0.05 (0.02, 0.13)	1.00 (0.97, 1.00)	1.00 (0.44, 1.00)	0.68 (0.61, 0.74)	~ (0.73, 265)*	0.95 (0.90, 1.01)[*]	3.1 (2.5, 3.8)
Nine	1	65	0	130	0.02 (0.00, 0.08)	1.00 (0.97, 1.00)	1.00 (0.21, 1.00)	0.67 (0.60, 0.73)	~ (0.24, 142)*	0.98 (0.94, 1.02)[*]	3.0 (2.5, 3.7)

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; ~, infinity

Note. Clinical examination results: strain mechanism of injury; anterior shoulder pain; unable to sleep on affected side; HBB, PROM GHJ abduction, PROM external rotation (at 90° abduction), PROM internal rotation (at 90° abduction), cross-body adduction (in internal rotation) do not provoke typical symptoms; negative Hawkins-Kennedy test. Variables were selected based upon association with an 80% positive anaesthetic response (PAR) ($p \le 0.200$)

Total cell counts are less than 196 for some variables due to missing data

*0.5 added to cells to estimate confidence intervals

The three variables identified in the clinical prediction model however, could not rule-out a PAR to SAB diagnostic block, with highest sensitivity of only 40% (one of three findings present). Possible explanations for the low sensitivity include the heterogeneity of subacromial pain and pathology in primary care populations, and the relatively low prevalence of an 80% PAR (34%). Structures that occupy the subacromial region including the SAB and components of the rotator cuff, cross the anatomical boundaries that were arbitrarily set for anterior shoulder pain (Precerutti, Garioni, Madonia, & Draghi, 2010). Thus lesions of the SAB or the rotator cuff, in the absence of anterior shoulder pain may still report relief from subacromial injections of local anaesthetic. Similarly, subacromial pain is known to result from mechanisms other than 'strain' including trauma, repetitive activity or insidious onset, and also as a result of inflammatory disease (Petersson, 1986).

Diagnostic Accuracy of Combinations of Clinical Examination Findings

Using combinations of the nine clinical variables however (Table 5.5), the ability to rule-out an 80% PAR (sensitivity) improved to 100% when a participant did not report at least one variable present. A PAR could also be ruled-out with a high level of confidence if a least two variables were not present (sensitivity 0.95; 95% CI 0.87, 0.98 and -LR 0.21; 95% CI 0.07, 0.62). Specificity of a PAR also increased with increasing numbers of positive tests, however there was a trade-off with decreasing numbers of participants satisfying the criteria that included higher numbers of positive tests. When six findings were positive and participants were almost seven times more likely to report a PAR to SAB diagnostic block (specificity 0.97), and when seven, eight or nine clinical tests were positive, specificity increased to 100%. When more expensive or invasive investigations or interventions are considered, higher diagnostic certainty would be achieved with higher numbers of positive clinical findings. In cases where clinical findings present diagnostic uncertainty (three, four of five positive clinical findings), a diagnostic injection of local anaesthetic into the subacromial region may be required to confirm the diagnosis. This is a simple and inexpensive diagnostic procedure when performed 'blind' in primary care with low associated risks, and in competent hands, injection accuracy approaches that of guided procedures (Rutten, Maresch, Jager, & Malefijt, 2007).

A limitation of note is that it cannot be precisely determined which structures were anaesthetized as no contrast agent was used during the diagnostic block procedure. However, recognized procedures were followed in the test protocol and the infiltration

of the SAB was confirmed in all cases. It is possible that structurally compromised portions of the rotator cuff may have been infiltrated with anaesthetic, and whether specific subgroups of subacromial pathology for which management decisions may be altered such as full thickness rotator cuff tears can be identified from clinical examination findings will be the subject of ongoing analysis. Further research is required to evaluate clinically meaningful anaesthetic response criteria and to assess the false-positive (placebo) rate of anaesthetic blocks around the shoulder.

Conclusion

In conclusion, the use of combinations of several clinical examination findings enables the clinician to select cut-points for numbers of clinical tests with levels of diagnostic accuracy that are compatible with the clinical objective. An 80% PAR could be ruled out if at least one of the nine clinical findings was not present. The clinical prediction model (three positive tests), and the presence of six or more of the nine clinical features enabled accurate identification of those participants likely to report an 80% PAR to SAB diagnostic block supporting the diagnosis of painful bursal conditions. For patients who fit these criteria, this provides confidence in the application of more expensive or invasive investigation or treatment interventions. However in the majority of cases, additional diagnostic tests such as diagnostic injections of local anaesthetic or diagnostic imaging may be required to confirm a subacromial pain source defined by an 80% anaesthetic response to subacromial bursa diagnostic block.

5.2. ADDED VALUE OF IMAGING FINDINGS FOR PREDICTING A POSITIVE RESPONSE TO GUIDED SUBACROMIAL BURSA DIAGNOSTIC BLOCK

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Taylor, S. The added value of imaging findings for predicting a positive response to guided subacromial bursa diagnostic block.

Results presented in the following manuscript have been formatted for journal submission upon completion of the PhD.

Abstract

Background: Clinical predictors of a PAR to SAB diagnostic block were identified in the previous section (strain injury, anterior shoulder pain and absence of symptom reproduction with passive external rotation). In clinical practice imaging is often used to assist in the diagnosis of subacromial disorders, however whether imaging findings improve the ability to identify those likely to report a PAR, indicating a SAB pain source, is unknown.

Objectives: To evaluate the impact on diagnostic accuracy of the addition of imaging findings to clinical examination findings for predicting a positive anaesthetic response (PAR) to subacromial bursa (SAB) diagnostic block.

Methods: Consecutive patients with shoulder pain underwent a standardised clinical examination, shoulder x-ray series and diagnostic ultrasound. Results were compared with the response to a diagnostic block of xylocaine™ injected into the SAB under ultrasound guidance. Multivariate regression analysis was used to develop prediction models for imaging findings with the strongest predictive ability for a positive anaesthetic response (PAR) (≥80% post-injection reduction in pain intensity). These variables were combined with clinical examination predictors and diagnostic accuracy statistics were calculated.

Results: A PAR was reported by 34% of participants. The strongest imaging predictors of a PAR were evidence of supraspinatus calcification and a full-thickness supraspinatus tear (FTT) on ultrasound. When only one (specificity 52%) or two (specificity 86%) clinical examination tests were positive, ultrasound findings of supraspinatus pathology (AOR 3.1) improved specificity to 99% (FTT) and 98%

(calcification) respectively (95% CI: 0.95, 1.00) with improvements in post-test probability to 78% and 80% respectively. Sensitivity for all models ranged from 3% to 40%.

Conclusion: Ultrasound findings of supraspinatus pathology improved the ability to rule-in a PAR when fewer clinical examination tests were positive. Additional diagnostic investigations may be required for those who do not fit the prediction model criteria.

Introduction

Different shoulder conditions frequently exhibit similar clinical characteristics and this increases the complexity of the diagnostic process in primary care, resulting in delayed diagnosis, protracted treatment courses and poor outcomes. Subacromial disorders are the most commonly reported shoulder condition, accounting for 50% to 70% of shoulder conditions seen in primary care (Cadogan, Laslett, Hing, McNair, & Coates, 2011; Chard et al., 1990; D. A. W. M van der Windt et al., 1995). Subacromial pain encompasses a range of disorders including subacromial bursitis, rotator cuff tendinopathy, rotator cuff tear and the clinical diagnosis of 'subacromial impingement' (Cyriax, 1982; Neer, 1983).

Diagnostic injection of local anaesthetic into the subacromial region is considered the reference standard test for identification of subacromial pain (Neer, 1983). Clinical examination features of strain mechanism of injury, anterior shoulder pain and absence of pain with external rotation at 90° abduction were the strongest clinical predictors of PAR (100% specificity when all three were positive) identified in the previous section. Combinations of nine clinical variables also demonstrated 100% sensitivity (at least one finding present) and 97% specificity (six or more findings present) for a PAR. However, when less than five of these findings were present, it was not possible to identify those who were likely to report a PAR (specificity 0.03 to 0.73), and when less than three findings were present it was not possible to identify those who were unlikely to report a PAR (sensitivity 0.02 to 0.77). In clinical terms this made it difficult to rule-out a likely subacromial pain source at the 80% pain relief standard, and it reduced confidence in identifying those who were likely to have a predominant subacromial pain source, or those for whom targeted pain relief interventions (e.g. corticosteroid injection) may be appropriate in the majority of cases.

Diagnostic imaging such as x-ray and ultrasound scans are being increasingly used to aid in the diagnostic process, adding further to the increasing cost of health care and resource utilisation (Awerbuch, 2008). Studies suggest the increasing use of diagnostic imaging for shoulder pathology may be related to low levels of practitioner confidence in the clinical diagnosis (Awerbuch, 2008; Johal et al., 2008). While imaging may provide evidence of pathological tissue changes, the high prevalence of asymptomatic pathology identified on imaging, particularly in ageing populations (Milgrom et al., 1995) can complicate the interpretation of imaging results with respect to symptomatic pathology. In the absence of clear guidelines for the use of diagnostic imaging, referral practices for investigation of shoulder pain at primary care level are inconsistent and often result in unnecessary referrals to specialist levels of care (Johal et al., 2008).

Little attention has been paid to the relationship between imaging findings and symptoms in those with shoulder pain. Our previous results have shown the prevalence of SAB pathology and rotator cuff pathology in a primary care cohort to be 31% and 50% respectively, yet only 34% of the cohort reported a PAR (Cadogan, Laslett, Hing, McNair, & Coates, 2011). Despite the similar prevalence of SAB pathology on ultrasound (31%) and PAR to SAB diagnostic block (34%), there was no clear association between SAB pathology or dynamic bursal bunching on ultrasound and a PAR to SAB diagnostic block (p>0.10) (Appendix 9, p276) suggesting those who reported a PAR did not demonstrate pathological SAB changes on ultrasound in many cases. In the clinical setting, it is unknown whether identification of subacromial pathology on imaging would alter decisions regarding treatment for subacromial pain. Knowledge of the relative additional diagnostic value of imaging results for identifying subacromial pain would also assist the clinician in making decisions regarding the need for referral for imaging investigations.

The aim of this study was to identify the strongest diagnostic imaging predictors of a PAR to SAB diagnostic block, and to evaluate to what extent these findings may alter the diagnostic accuracy for a PAR when combined with clinical examination findings.

Methods

Design, Sampling and Recruitment

The study design, ethical approval, sampling, recruitment and clinical examination procedures (index tests) are the same as those reported in the previous section.

Imaging Procedures and Subacromial Bursa Diagnostic Block

The diagnostic imaging procedures (x-ray and ultrasound scan), SAB diagnostic block procedures and calculation of change in pain intensity were performed as described in Chapter 4 (p107). Radiological diagnostic criteria were presented in Table 4.1 (p106).

Statistical Analysis

The association of each imaging variable with a PAR to SAB diagnostic block was assessed with Fisher's exact test, using Statistical Package for the Social Sciences (SPSS) version 17.0 (IBM® Corporation 2010). Two separate multivariate prediction models were then considered. The first (model 1) included clinical examination variables identified in the previous section. The second model was derived from model 1 by the addition of imaging variables to clinical examination variables (model 1), and the process of stepwise backward variable elimination was repeated to derive the strongest combination of clinical examination and imaging predictors (model 2). A minimum of five outcome events were present per predictor variable in the multivariate analyses (Rawlings, Rae, & Graubard, 1982; Wasson, Sox, Neff, & Goldman, 1985). Multiple regression analysis was carried out using "R", a language and environment for statistical computing (R Development Core Team, 2010).

The accuracy and discriminatory ability of these models was assessed by calculating sensitivity, specificity, predictive values, positive likelihood ratios (+LR) and negative likelihood ratios (-LR) and 95% confidence intervals (CI) for combinations of clinical examination and imaging findings using Confidence Interval Analysis software (Bryant, 2000). Post-test odds {pre-test odds x +LR}, and post-test probability {post-test odds/[1+post-test odds]} of a PAR to SAB diagnostic block were calculated for each combination of clinical and imaging results (Schwartz, 2002). The goodness of fit for models was assessed using the Hosmer-Lemeshow test (Hosmer & Lemeshow, 2000). To assess the incremental value of adding imaging variables, ROC

analyses were carried out using predicted probabilities from model 1 and model 2 and the AUC was compared for each model.

Results

Participant demographics are the same as those presented in Table 5.1 (p132). Results for recruitment, procedure completion rates and results were presented in Figure 5.1 (p133).

Subacromial Diagnostic Block

Descriptive results for the SAB diagnostic block were presented in Chapter 4 (p116). A PAR (≥80% pain relief) was reported by 66 of the 196 (34%) cases following the SAB injection. A graphical summary of frequency of major pathology categories in the PAR and NAR groups is presented in Figure 5.2. Frequency distributions of specific imaging findings for the PAR and NAR groups are presented in Table 5.6.
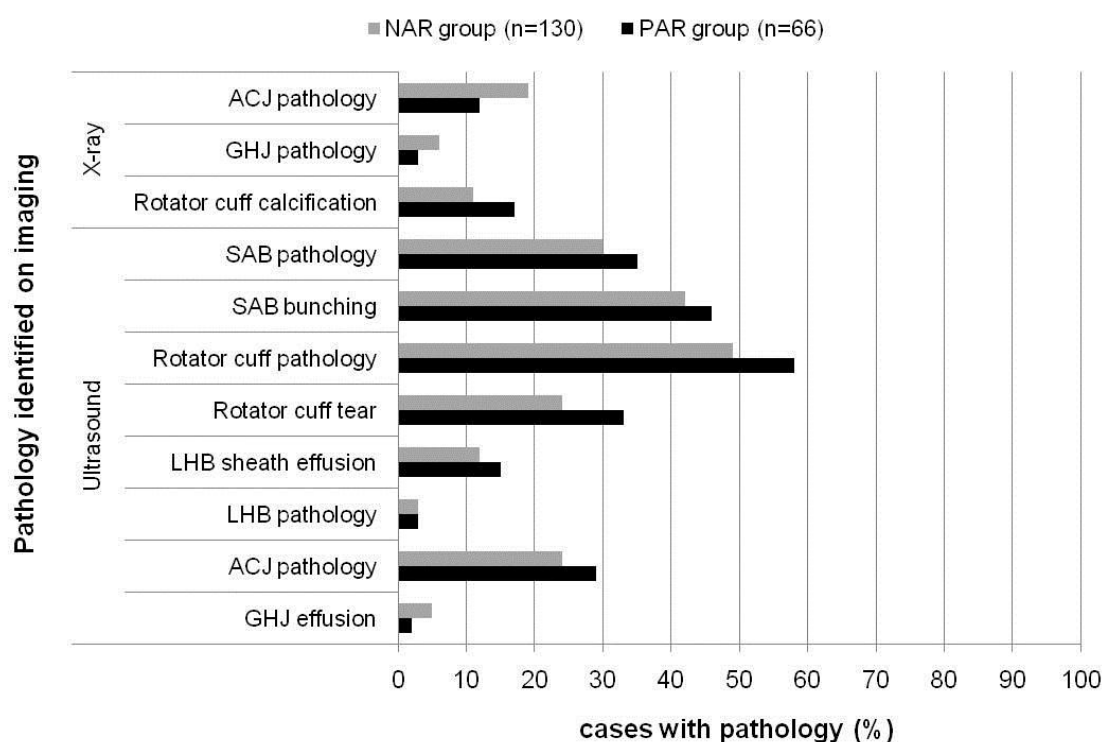
Figure 5.2. Major pathologies in subacromial bursa diagnostic block PAR and NAR groups. Graph showing differences in frequency distribution of imaged pathology in the PAR and NAR groups. There were no significant differences between the frequency of major pathological categories in the PAR or NAR groups (*p*>0.05). NAR, negative anaesthetic response; PAR, positive anaesthetic response; ACJ, acromioclavicular joint; GHJ, glenohumeral joint; SAB, subacromial bursa; LHB, long head of biceps.

Table 5.6. *Distribution of Imaged Pathology in SAB Diagnostic Block PAR and NAR Groups.*

Diagnostic test results	Total identified	% in PAR group with pathology	% in NAR group with pathology
	(N)	(n=66)	(n=130)
X-ray			
ACJ pathology	32	12	19
ACJ arthropathy	24	9	14
ACJ osteolysis	7	2	5
GHJ pathology	10	3	6
Rotator cuff calcification	25	17	11
supraspinatus	16	14	5*
infraspinatus	7	3	4
subscapularis	6	0	5
Ultrasound			
SAB pathology[a]	62	35	30
Bursal bunching			
acromion	81	46	42
CAL[b]	51	54	55
Rotator cuff tear	53	33	24
supraspinatus PTT	12	3	8
supraspinatus FTT	10	11	2**
infraspinatus PTT	1	0	1
Infraspinatus FTT	1	0	1
subscapularis PTT	4	2	2
subscapularis FTT	1	0	1
Rotator cuff tendinosis	29	14	15
supraspinatus	27	14	14
infraspinatus	1	0	1
subscapularis	4	0	3
Rotator cuff calcification	48	26	24
supraspinatus	33	24	13*
infraspinatus	9	5	5
subscapularis	20	8	12
LHB tear or tendinosis	6	3	3
LHB tendon sheath effusion	26	15	12
ACJ pathology	50	29	24
GHJ effusion	7	2	5

Abbreviations: PAR, positive anaesthetic response; NAR, negative anaesthetic response; ACJ, acromioclavicular joint; GHJ, glenohumeral joint; SAB, subacromial bursa; CAL, coracoacromial ligament; PTT, partial thickness tear; FTT, full thickness tear; LHB, long head of biceps.

Note. Pathology subgroup totals may exceed composite pathology totals due to some cases identified in which multiple pathologies were present.

[a]SAB pathology included: thickening ≥2mm, calcification, bursal fluid or effusion.

[b] bunching under the CAL only assessed in 93 cases. (PAR n=26; NAR n=67)

*$p<0.05$; ** $p<0.01$; *** $p≤0.001$.

Prediction Models

Imaging variables associated with a PAR to SAB diagnostic block were supraspinatus calcification on x-ray ($p=0.054$), supraspinatus calcification on ultrasound ($p=0.068$) and a full-thickness supraspinatus tear (FTT) on ultrasound ($p=0.033$). Due to covariance of supraspinatus calcification on both x-ray and ultrasound, both variables

were unable to be included in the regression analysis. The x-ray variable was excluded as in some clinical situations an ultrasound scan is requested without concurrent x-ray investigations and two of the three variables associated with a PAR were identified on ultrasound. Due to the low prevalence of FTT, the two ultrasound variables (supraspinatus calcification and FTT) were combined into a single, composite variable with a positive "imaging" case defined as those in whom either one of these variables was present. Contingency cell counts and diagnostic statistics for imaging variables associated with a PAR are presented in Table 5.7.

Table 5.8 presents the clinical examination variables included in multivariate prediction models after backward stepwise variable elimination to derive model 1 (clinical examination). Clinical variables retained in prediction model 1 were pain location (anterior), mechanism of injury (strain) and the absence of symptom provocation during passive ROM external rotation (at 90° abduction). Imaging variables were then added to model 1 and further backward stepwise variable elimination was conducted to derive model 2 (clinical examination and imaging). The composite imaging variable was retained in model 2 (AOR 3.1 (95% CI 1.5, 6.6)).

Evaluation and Accuracy of the Models

Hosmer-Lemeshow statistics indicate the goodness of fit of both models was adequate (clinical examination model χ^2_6=3.24, p= .778; clinical and imaging model χ^2_6=2.46, p= .872). Diagnostic accuracy and post-test probability for individual and combinations of clinical examination and imaging variables for a PAR to SAB diagnostic block are presented in Table 5.9. Overall sensitivity was low and ranged from 0.03 (all three clinical tests positive and US evidence of supraspinatus calcification or FTT) to 0.40 (any one of three clinical tests positive). Specificity ranged from 0.52 (any one of three clinical tests positive) to 1.00 (all three clinical tests positive irrespective of imaging findings, and all three clinical tests positive with US evidence of supraspinatus FTT). Positive predictive values ranged from 0.21 to 1.00 (all three clinical tests positive) and negative predictive values (NPV) ranged from 0.67 to 0.72 (any one clinical test positive and supraspinatus calcification or FTT on US). Highest +LR (7.76) and post-test probability (80%) for PAR to SAB diagnostic block was observed when any two clinical tests were positive and supraspinatus calcification was identified on US. The lowest -LR was 0.77 (any two clinical tests positive irrespective of imaging findings).

Table 5.7. *Diagnostic Accuracy of Imaging Variables Associated with a Positive Response to Subacromial Bursa Diagnostic Block*

Imaging variables	Cell counts				Diagnostic accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)
X-Ray: supraspinatus calcification	9	56	7	123	0.14 (0.08, 0.24)	0.95 (0.89, 0.97)	0.56 (0.33,0.77)	0.69 (0.62,0.75)	2.57 (1.03, 6.38)	0.91 (0.80, 1.00)	2.82 (1.00, 7.97)
US: Supraspinatus calcification	16	50	17	113	0.24 (0.16, 0.36)	0.87 (0.80, 0.92)	0.49 (0.33,0.65)	0.69 (0.62,0.76)	1.85 (1.00, 3.38)	0.87 (0.73, 1.00)	2.13 (1.00, 4.55)
US: Supraspinatus FTT	7	59	3	127	0.11 (0.05, 0.20)	0.98 (0.93, 0.99)	0.70 (0.40,0.89)	0.68 (0.61,0.75)	4.60 (1.33, 15.89)	0.92 (0.81, 0.98)	5.02* (1.25, 20.11)

Abbreviations. TP, true positive; FN, false negative; FP, false positive; TN, true negative; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; US, ultrasound; SSp, supraspinatus; FTT, full thickness tear
*$p<0.05$.

Table 5.8. *Clinical Examination and Imaging Prediction Models for a Positive Response to Subacromial Bursa Diagnostic Block*

Variables	Clinical examination (model 1) n=193*		Clinical examination and imaging (model 2) n=193*	
	AOR	95% CI	AOR	95% CI
Strain	2.3	1.2, 4.4	2.2	1.2, 4.3
Anterior pain	2.3	1.2, 4.5	2.4	1.2, 4.7
PROM ER90- asymptomatic	3.9	1.9, 8.0	4.1	2.0, 8.9
US: SSp calcification or FTT			3.1	1.5, 6.6

Abbreviations. AOR, adjusted odds ratio; CI, confidence interval; PROM ER90, passive range of motion external rotation (at 90° abduction); US, diagnostic ultrasound; SSp, supraspinatus; FTT, full thickness tear
*three of the 196 participants were excluded from both models due to missing covariate values

Table 5.9. *Diagnostic Accuracy of Clinical Examination and Imaging Prediction Models for a Positive Response to Subacromial Bursa Diagnostic Block*

Clinical Prediction Model[a]	Cell counts				Diagnostic accuracy					
					Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)
	TP	FN	FP	TN						
CPM 1	26	39	62	66	0.40 (0.29, 0.52)	0.52 (0.43, 0.60)	0.30 (0.21, 0.40)	0.63 (0.53, 0.72)	0.8 (0.6, 1.2)	1.16 (0.89, 1.57)
CPM1 and imaging:										
a) No SSp calc or FTT	15	51	55	75	0.23 (0.14, 0.34)	0.58 (0.49, 0.66)	0.21 (0.13, 0.32)	0.60 (0.51, 0.68)	0.5 (0.3, 0.9)	1.34 (1.09, 0.63)
b) SSp calc	4	62	7	123	0.06 (0.02, 0.15)	0.95 (0.89, 0.97)	0.36 (0.15, 0.65)	0.67 (0.59, 0.73)	1.1 (0.4, 3.5)	0.99 (0.90, 1.07)
c) SSp FTT	7	59	2	128	0.11 (0.05, 0.20)	0.99 (0.95, 1.00)	0.78 (0.45, 0.94)	0.68 (0.62, 0.75)	6.9 (1.7, 28.6)	0.91 (0.81, 0.97)
d) either SSp calc or FTT	11	55	9	121	0.17 (0.10, 0.27)	0.93 (0.87, 0.96)	0.55 (0.34, 0.74)	0.69 (0.62, 0.75)	2.4 (1.1, 5.4)	0.90 (0.78, 0.99)
CPM 2	22	43	18	110	0.34 (0.24, 0.46)	0.86 (0.79, 0.91)	0.55 (0.40, 0.69)	0.72 (0.64, 0.78)	2.4 (1.4, 4.1)	0.77 (0.62, 0.91)
CPM 2 and imaging:										
a) No SSp calc or FTT	15	51	16	112	0.23 (0.14, 0.34)	0.88 (0.81, 0.92)	0.48 (0.32, 0.65)	0.69 (0.61, 0.75)	1.8 (1.0, 3.4)	0.88 (0.75, 1.01)
b) SSp calc	8	58	2	126	0.12 (0.06, 0.22)	0.98 (0.95, 1.00)	0.80 (0.49, 0.94)	0.69 (0.62, 0.75)	7.8 (1.9, 31.7)	0.89 (0.79, 0.96)
c) SSp FTT	0	66	0	128	-	-	-	-	-	-
d) either SSp calc or FTT	8	58	2	126	0.12 (0.06, 0.22)	0.98 (0.95, 1.00)	0.80 (0.49, 0.94)	0.69 (0.62, 0.75)	7.8 (1.9, 31.7)	0.89 (0.79, 0.96)
CPM 3	6	59	0	128	0.09 (0.04, 0.19)	1.00 (0.97, 1.00)	1.00 (0.61, 1.00)	0.68 (0.62, 0.75)	~ (1.5, 444.0)[†]	0.91 (0.84, 0.98)[†]
CPM 3 and imaging:										
a) No SSp calc or FTT	4	61	0	130	0.06 (0.02, 0.15)	1.00 (0.97, 1.00)	1.00 (0.51, 1.00)	0.68 (0.61, 0.74)	~ (1.0, 327.0)[†]	0.94 (0.88, 1.00)[†]
b) SSp calc	2	63	0	130	0.03 (0.01, 0.11)	1.00 (0.97, 1.00)	1.00 (0.34, 1.00)	0.67 (0.61, 0.74)	~ (0.5, 204.0)[†]	0.97 (0.92, 1.01)[†]
c) SSp FTT	0	65	0	130	-	-	-	-	-	-
d) either SSp calc or FTT	2	63	0	130	0.03 (0.01, 0.11)	1.00 (0.97, 1.00)	1.00 (0.34, 1.00)	0.67 (0.61, 0.74)	~ (0.5, 204.0)[†]	0.97 (0.92, 1.01)[†]

Abbreviations. TP, true positive; FN, false negative; FP, false positive; TN, true negative; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; CPM, clinical prediction model; ~, infinity; SSp, supraspinatus; calc, calcification; FTT, full thickness tear; US, ultrasound

[a]Clinical prediction model based on the following examination tests: strain mechanism of injury, anterior shoulder pain and absence of symptom reproduction with passive range of motion external rotation (at 90° abduction): CPM 1, any one of three clinical tests positive; CPM 2, any two of three clinical tests positive; CPM 3, all three clinical tests positive. All imaging relates to pathology identified on diagnostic ultrasound scan

[b]post-test probability for PAR to SAB diagnostic block. Calculated using the positive likelihood ratio assuming 34% prevalence of positive anaesthetic response to diagnostic block based on prevalence in this study

- no cases identified, invalid calculation

[†]represents estimated post-test probability and 95% CI based on +LR of infinity

Additional Value of Imaging

The AUCs for predicted probabilities of models 1 and 2 were respectively 0.693 (95%CI: 0.609, 0.773; $p<0.001$) and 0.731 (95%CI: 0.651, 0.811; $p<0.001$), the shift to the left of the ROC curve for model 2 indicating improvement in discriminatory ability for participants with PAR with the inclusion of the ultrasound imaging variable Figure 5.3. Identification of a supraspinatus FTT on ultrasound improved specificity for a PAR from 0.52 to 0.99 when only one clinical examination finding was present (post-test probability 78%, (Figure 5.4), and improved specificity for a PAR from 0.86 to 0.98 when two clinical findings were present (post-test probability 80%) (Figure 5.4). The addition of imaging findings resulted in a reduction in sensitivity for a PAR for all models.
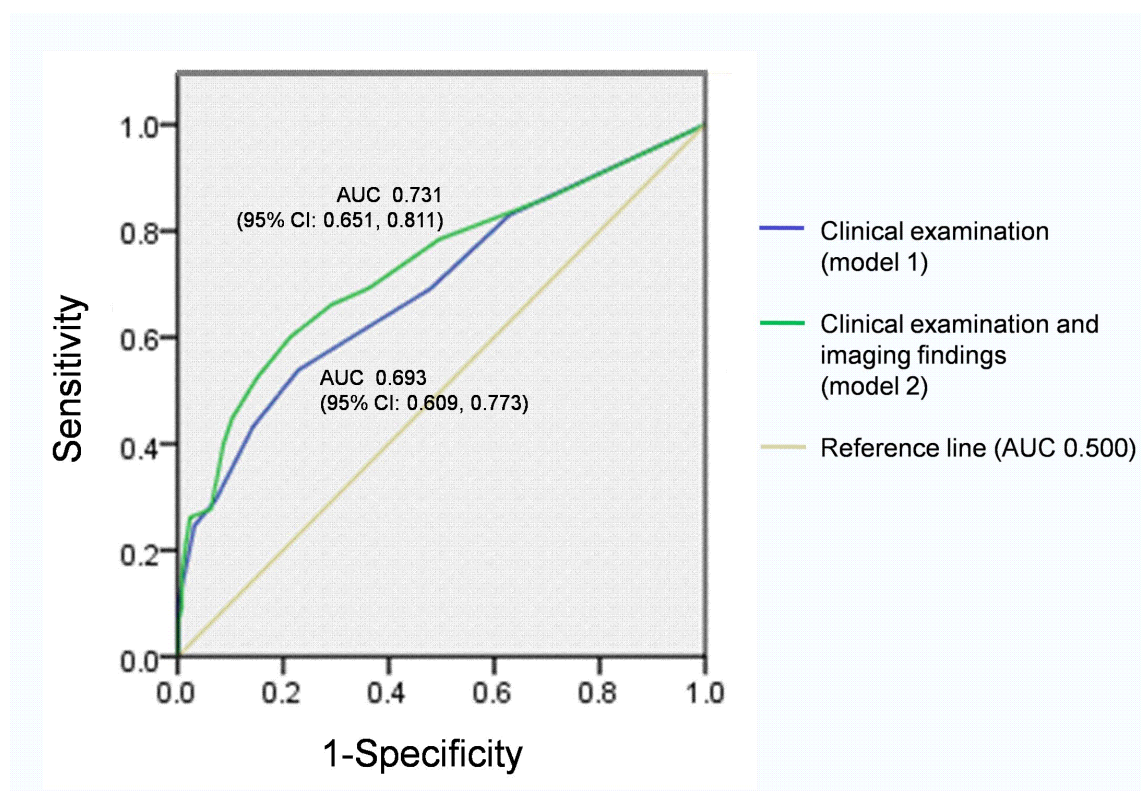
Figure 5.3. Receiver operator curve for clinical and imaging prediction models. *Abbreviation*: AUC, area under curve.
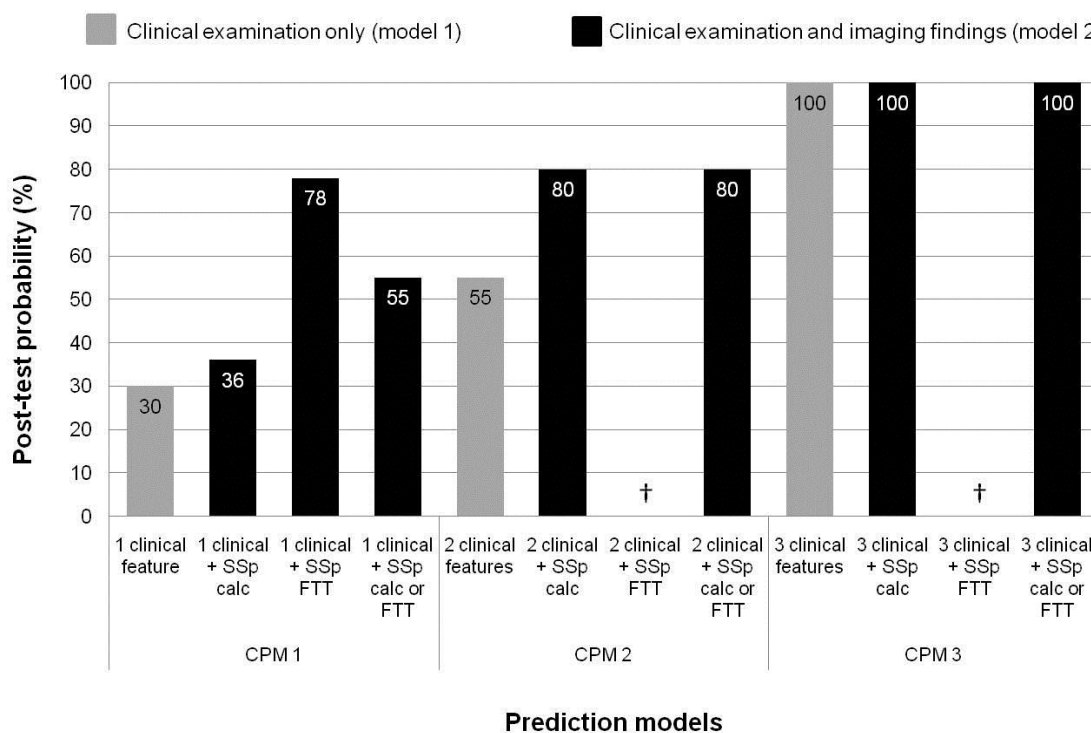
Figure 5.4. Change in post-test probability for a PAR following SAB diagnostic block with addition of imaging findings to clinical examination findings. Graph showing change in post-test probability (%) when model 2 (imaging) variables are added to model 1 (clinical examination variables). SSp, supraspinatus; calc, calcification; FTT, full-thickness tear; CPM, clinical prediction model; †, invalid calculation, no cases identified.

Discussion

This is the first report of a prospective cohort of patients with shoulder pain recruited from primary care in which imaging predictors of a PAR to SAB diagnostic block were derived, and the relative value of imaging findings for predicting a PAR when combined with clinical examination findings was evaluated. Ultrasound imaging findings of supraspinatus pathology (calcification or full thickness tear) were predictive of a PAR. Prediction models are presented that may aid clinicians in decisions regarding the relative value of imaging findings for identifying those likely to report an 80% PAR following a SAB diagnostic block.

Prediction models may provide the clinician with an adjunct to clinical reasoning during the diagnostic processes where the ability to rule-in a PAR to SAB diagnostic block would assist differential diagnosis of subacromial pain from other shoulder disorders. They may also inform decision making regarding selection of treatment interventions targeted at commonly reported causes of subacromial pain including scapula dyskinesis (Burkhart et al., 2003; B. W. Kibler, 1998) and humeral head stability (Warner et al., 1990) and may provide an indication of those who may benefit

from more targeted, expensive or invasive treatment interventions such as barbotage (Comfort & Arafiles, 1978) corticosteroid injection or surgery. Identification of subacromial pain also provides rationale for avoiding those interventions known to increase subacromial pressure such as positions of glenohumeral internal rotation (Werner, Blumenthal, Curt, & Gerber, 2006). The use of tests with high levels of specificity (low false positive rate) reduces the risk of adverse events for invasive procedures, reduces costs of inappropriate procedures, and may improve patient outcomes by selecting appropriate interventions for specific disorders (Childs & Cleland, 2006). The ability to rule-out subacromial pain (high levels of sensitivity) would enable the clinician to identify those patients for whom such interventions are not appropriate, and may enhance the diagnostic process by refocusing the evaluation on other potential sources of symptoms.

Diagnostic Imaging Predictor Variables

Clinical examination predictors of a PAR to SAB diagnostic block are discussed in more detail in the previous section. This is the first known report of imaging predictors of a positive response to SAB diagnostic block.

A full thickness supraspinatus tear identified on ultrasound was the strongest independent predictor of a PAR to SAB diagnostic block and those with a supraspinatus FTT were five times more likely to report a PAR at the 80% pain relief level than those without a FTT (OR 5.02). Communication of anaesthetic with supraspinatus through the disruption to the SAB-rotator cuff interface is the likely explanation for this result. The prevalence of full thickness rotator cuff tears in asymptomatic participants is known to increase with age (Milgrom et al., 1995). In this sample of participants suffering a current episode of shoulder pain, age was not associated with anaesthetic response and, when present, a FTT demonstrated 98% specificity for a PAR. These results suggest that although the prevalence of asymptomatic tears may increase with age, in patients suffering a current episode of shoulder pain these lesions are of symptomatic significance, irrespective of age.

While prognostic predictors for patients suffering shoulder pain have been previously reported (Kuijpers et al., 2006; Macfarlane, Hunt, & Silman, 1998; D. A. W. M van der Windt et al., 1996), no other studies were found in which clinical or imaging predictors had been derived for the diagnosis of subacromial pain using injection of local anaesthetic as the reference standard test.

Prediction Models

Our results suggest that subacromial pain cannot be ruled out with any degree of certainty using either of the prediction models derived in this study. Sensitivity for all prediction models was low, the highest (40%) was observed for CPM1 in which any one clinical test was positive, irrespective of imaging findings, meaning many who may report a PAR were not identified using this criterion. The lowest -LR was 0.77 for CPM2 resulting in a moderate probability (28%) that someone who didn't have at least two positive tests would still report a PAR. A possible explanation for the low sensitivity is the heterogeneity of subacromial pain and pathology, and different histological features have previously been associated with different clinical presentations (Santavirta, Konttinen, Antti-Poika, & Nordstrom, 1992; Sarkar & Uhthoff, 1983). Further subgroup analysis of SAB pathology may add to this information.

The clinical implications of low sensitivity of these prediction models are potentially prolonged diagnostic processes that result in unnecessary investigations or referrals for assessment with associated costs to the health care system, and prolonged periods of pain and disability with a resulting decline in treatment outcomes. Additional diagnostic procedures such as a clinically administered diagnostic injection of local anaesthetic may be required to confirm the anaesthetic response. A diagnostic injection of local anaesthetic into the subacromial region, while used as the reference standard test in this study, has also been used in the clinical setting for many years to assist in confirming the diagnosis of subacromial pain when a reduction in post-injection pain intensity is reported (Cyriax, 1982; Neer, 1983; Ombregt, Bisschop, & ter Veer, 2003). This is a simple and inexpensive diagnostic procedure when performed 'blind' in primary care with low associated risks and, in competent hands, injection accuracy approaches that of guided procedures (Rutten et al., 2007). A subacromial diagnostic block provides an immediate indication as to whether the subacromial structures are symptomatic. It is also less expensive than further imaging investigations, specialist consultations or a course of treatment based upon a 'wait and see' approach. A lack of training or inability to perform these injections was one of the reasons commonly cited by general 'internists' for referral for specialist consultation resulting in many unnecessary referrals to secondary care services (Donohoe et al., 1999). A diagnostic injection of local anaesthetic into the subacromial space would appear to be a reasonable procedure for those whose clinical examination findings are not specific for a subacromial source of pain, and this may require further training of primary care

practitioners or extended scope of practice for physiotherapists to facilitate wider access to these diagnostic tests.

Added Value of Imaging Findings

The addition of imaging findings to the clinical examination prediction model resulted in improvements in diagnostic accuracy and post-test probability of a PAR. Our results demonstrated an increase in the ability of the predictive model to discriminate between those who demonstrated a PAR to SAB diagnostic block when the imaging findings were added according to the AUC. When imaging findings were added to clinical models in which only one (CPM1) or two (CPM2) clinical features were present specificity improved to 0.93 and 0.98 in the respective models. The report of either supraspinatus calcification or a FTT also increased the post-test probability of a PAR for both models from 30% and 55%, to 78% and 80% for CPM1 and CPM2 respectively, representing a clinically meaningful increase in probability of a PAR. The lower confidence limits for the post-test probabilities (42%, 43% and 47% respectively) did however represent only a moderate improvement over pre-test probability and some caution should be applied when interpreting these results.

Although the addition of imaging variables to clinical examination findings increased the specificity for a PAR, this came at the expense of sensitivity, which was reduced with the addition of imaging findings in all models. The lower sensitivities when imaging variables were included in the prediction models were primarily due to the low prevalence of imaging findings. The number of participants who satisfied the prediction model criteria that included imaging findings was reduced to between 1% (CPM3d) and 5% (CPM2d) which limits the number of patients to whom the models would apply in the clinical setting.

Limitations of the Study

Reports of anaesthetic response criteria following diagnostic blocks into peripheral joint and peri-articular structures are scarce and further research is required to evaluate clinically meaningful anaesthetic response criteria and to assess the false-positive (placebo) rate of anaesthetic blocks around the shoulder. It cannot be precisely determined which structures were anaesthetized in this study as no contrast was used during the diagnostic block procedure, and although infiltration of the SAB was confirmed in all cases, the rotator cuff may also have been exposed to anaesthetic (Kuhn & McGuigan, 2006). The anaesthetic response may thus reflect a secondary bursitis

and/or anaesthesia of the sensitive cuff which is in direct contact with the bursal space. Due to the relatively low prevalence of imaging findings, further testing on larger samples is required to confirm these results. We do not propose these prediction models be used as screening tests or clinical prediction rules in their current form due to low sensitivities, and the need for prospective validation (Laupacis, Sekar, & Stiell, 1997).

Conclusion

Ultrasound imaging findings of supraspinatus calcification or a full-thickness supraspinatus tear improved the ability to rule-in an 80% PAR following SAB diagnostic block when only one or two of the clinical examination predictors were identified. However, the low prevalence of these imaging findings means the clinical decision as to whether imaging findings are likely to be of value for identification of a PAR would need to be weighed against the availability and cost of the procedure. Diagnostic injection of local anaesthetic may be the most efficient and cost-effective method of identifying subacromial pain in primary care patients who do not fit these clinical prediction model criteria.

CHAPTER SIX

PREDICTORS OF RESPONSE TO ACROMIOCLAVICULAR JOINT DIAGNOSTIC BLOCK

Preface

This chapter relates to Specific Aims 2 and 4 of the thesis:

To estimate the diagnostic accuracy of clinical examination findings for identifying a predominant ACJ pain source defined by a positive response to diagnostic block.

To evaluate the added diagnostic value of imaging findings for predicting a positive response to ACJ diagnostic block.

This chapter presents two manuscripts reporting the diagnostic accuracy of clinical examination variables in predicting a positive response to ACJ diagnostic block, and the added diagnostic value of imaging findings in predicting an 80% ACJ PAR.

The first manuscript presents the diagnostic accuracy of a combination of nine clinical examination variables that were associated with a PAR following ACJ diagnostic block. In the second manuscript, the diagnostic accuracy of five of the variables that were most closely associated with a PAR was re-calculated and imaging findings were then added to the clinical examination variables. Diagnostic accuracy was then re-calculated and compared with accuracy of clinical examination features alone to assess the relative added value of imaging findings for predicting a 80% ACJ PAR.

Results for the added value of imaging findings for predicting a PAR are presented as a preliminary report only, as collection of data relating to specific ACJ pathology on ultrasound was not standardised to the same level as other ultrasound findings for reasons outlined in the discussion section of this manuscript. Results may be used as pilot data from which more investigation of the diagnostic value of ACJ pathological changes on ultrasound may be based.

The manuscripts in this chapter contain some repetition of methodology that has already been described elsewhere.

6.1. CLINICAL DIAGNOSIS OF A POSITIVE RESPONSE TO GUIDED ACROMIOCLAVICULAR JOINT DIAGNOSTIC BLOCK

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Taylor, S. Clinical diagnosis of a positive response to guided acromioclavicular joint diagnostic block *(under review)*.

Abstract

Background: Acromioclavicular joint (ACJ) pain is frequently seen in clinical practice for which specific interventions such as injection therapy are commonly used.

Question: The aim was to identify which clinical examination findings provide the highest levels of diagnostic accuracy for a positive response to imaging-guided ACJ diagnostic block.

Patients and methods: Consecutive patients with shoulder pain were recruited prospectively from primary health care clinics. Following a standardised clinical examination all participants received a fluoroscopically guided diagnostic block of 1% lidocaine hydrochloride (Xylocaine[TM]) into the ACJ. Diagnostic accuracy statistics were calculated for individual and combinations of clinical examination variables associated with a positive anaesthetic response (PAR) ($p \leq 0.200$) defined as 80% or more reduction in post-injection pain intensity during provocative clinical tests.

Results: Twenty two of 153 participants (14%) reported an 80% PAR. A repetitive mechanism of injury demonstrated highest specificity (0.90; 95% CI 0.84, 0.94) for a PAR and the absence of referred pain below the elbow demonstrated highest sensitivity (1.00: 95% CI 0.84, 1.00). Combinations of clinical examination variables improved specificity, which exceeded 80% when five or more tests were positive. Highest specificity was observed when eight positive tests were present (1.00; 95% CI 0.97, 1.00).

Conclusions: Combinations of history and physical examination tests enable identification of those patients likely to report an 80% PAR to ACJ diagnostic block when at least five clinical features were identified. This may inform diagnostic decision

making regarding the source of pain, guide clinical decisions regarding the use of targeted pain relief interventions such as corticosteroid injections, and provide rationale for the use of more expensive or invasive procedures.

Introduction

Shoulder pain is a common and disabling complaint and is frequently seen in primary care practice (Urwin et al., 1998). Disorders of the acromioclavicular joint (ACJ) are a common cause of shoulder pain affecting patients of all ages and levels of activity (Shaffer, 1999). While much has been written about traumatic ACJ instability (Deitch, 2004; Johansen, Grutter, McFarland, & Petersen, 2011; Mazzocca et al., 2007), ACJ pain may also be caused by a range of other pathologies including degenerative or post-traumatic arthropathy, inflammatory arthropathy, crystal arthropathy and osteolysis (A. P. Wright, MacLeod, & Talwalker, 2011). Identification of the ACJ as the primary source of pain is important to enable efficient application of appropriate treatment interventions, as well as to inform decisions regarding referral for further medical or imaging investigations or specialist consultation.

In clinical practice the diagnosis of shoulder pain begins with a clinical examination including patient history and physical examination. While several studies evaluated the diagnostic accuracy of physical examination tests for identifying ACJ disorders, the diagnostic value of aspects of patient history has not been previously reported. The active compression test is reported to be of diagnostic value for identifying ACJ pathology (O'Brien et al., 1998), however the x-ray and magnetic resonance imaging investigations used as reference standard procedures in this study do not take into account whether the pathology observed on imaging is the likely source of symptoms. To date the high levels of sensitivity (93%) and specificity (96%) reported in this study have not been independently verified.

The diagnostic accuracy of physical examination tests using a positive response to intra-articular injection of local anaesthetic into the ACJ as the reference standard test has also been investigated (Chronopoulos et al., 2004; Van Riet & Bell, 2011; Walton et al., 2004). Only one of these studies reported the use of imaging guidance to ensure accuracy of needle placement within the ACJ (Walton et al., 2004). Injections into the ACJ performed without the use of image intensification may be misplaced in up to 60% of cases (Bisbinas et al., 2006), casting doubt upon the validity of the reference standard

procedure in the other studies. These studies all recruited patients from specialist levels of care, and no studies were found involving primary care patient populations.

Diagnostic injections of local anaesthetic are often used in clinical practice to differentiate, or confirm a pathoanatomic diagnosis where a post-injection reduction in pain intensity is indicative of a positive result (Larson et al., 1996). However, imaging guidance is rarely available to ensure accurate placement of intra-articular ACJ injections in this setting, and diagnostic injections may fall outside the scope of practice for some practitioners. In order to facilitate efficient implementation of appropriate treatment interventions, or referral for further investigation or assessment of ACJ pain, clinical examination findings with high levels of diagnostic accuracy for a positive anaesthetic response to diagnostic injection are required. The purpose of this study was to identify clinical examination variables with the highest levels of diagnostic accuracy for a positive anaesthetic response to guided acromioclavicular joint diagnostic block.

Methods

Study Design and Setting

This study formed part of a wider prospective, blinded diagnostic accuracy study in which clinical examination and imaging variables (index tests) were compared with results of guided diagnostic injection of local anaesthetic (reference standard) into the subacromial bursa (SAB), acromioclavicular joint (ACJ) and glenohumeral joint (GHJ) (Cadogan, Laslett, Hing, McNair, & Coates, 2011). Participants were recruited from community-based medical and physiotherapy practices across Christchurch, New Zealand.

Recruitment and Sampling

Consecutive patients over the age of 18 years, presenting to their primary care practitioner (general practitioner or physiotherapist) for the first time with a new episode of shoulder pain (Figure 4.1) and with the ability to follow verbal instructions were eligible for inclusion in the study. Exclusion criteria were known fractures or dislocations around the shoulder complex, referred pain from the cervical spine, sensory or motor deficit involving the upper limb, previous surgery to the shoulder or cervical spine, or contraindications to imaging or injection procedures.

Sample size was estimated using methods for estimates for diagnostic accuracy studies described by Flahault et al (Flahault et al., 2005). The minimum acceptable

lower confidence limit was set at 0.75 and expected sensitivity/specificity were both set at 0.90. A subgroup analysis after the first 100 cases indicated the prevalence of ACJ pain was less than expected and sample size adjusted to maintain precision of diagnostic estimates.

Clinical Examination

All participants completed self-report questionnaires including SF-8[TM] health survey (Ware et al., 2001), Shoulder Pain and Disability Index (SPADI) (Roach et al., 1991) and Fear Avoidance Beliefs Questionnaire (FABQ) (Gordon Waddell et al., 1993). All participants recorded a standardised history including medical and family history, smoking history, a pain drawing, details of the current episode (duration of symptoms and mechanism of onset), details of past history of shoulder pain, occupational, sporting and recreational activities. A full list of clinical examination variables is presented in Appendix 7 (p273). All clinical examinations were conducted by a musculoskeletal physiotherapist with 20 years experience (AC).

The physical examination consisted of active range of motion (ROM) of the cervical spine (Maitland, 1986), inspection for swelling or muscle atrophy, recording the presence of a painful arc of motion during abduction (Kessel & Watson, 1977), recording of symptom responses associated with arm elevation (flexion), scapuloclavicular tests (elevation/depression and protraction/retraction) (Laslett, 1996), passive ROM glenohumeral abduction, external rotation performed at 0° abduction and internal and external rotation performed at 90° of abduction (Cadogan, Laslett, et al., 2011a), cross-body adduction performed in both internal and external rotation, resisted muscle tests (abduction, external and internal rotation), orthopaedic tests selected according to evidence for reported diagnostic accuracy (Hegedus et al., 2008) and performed as described by the original authors; Hawkins-Kennedy test (Hawkins & Kennedy, 1980), empty can test (F. Jobe & Moynes, 1982), active compression (O'Brien's) test (O'Brien et al., 1998), Speed's test (Gill et al., 2007), and pain responses to palpation of the ACJ.

Symptom responses were recorded during all ROM and resisted tests according to whether or not they reproduced typical pain. During the physical examination, those tests provocative of typical pain were identified for use in pre- and post-injection testing. Indeterminate results of clinical examination tests were recorded and coded as missing data. Following the clinical examination, all participants received a standardised shoulder x-ray series, diagnostic ultrasound scan, and subacromial bursa

diagnostic block as part of the larger diagnostic accuracy study (Figure 6.1) (Cadogan, Laslett, Hing, McNair, & Coates, 2011).
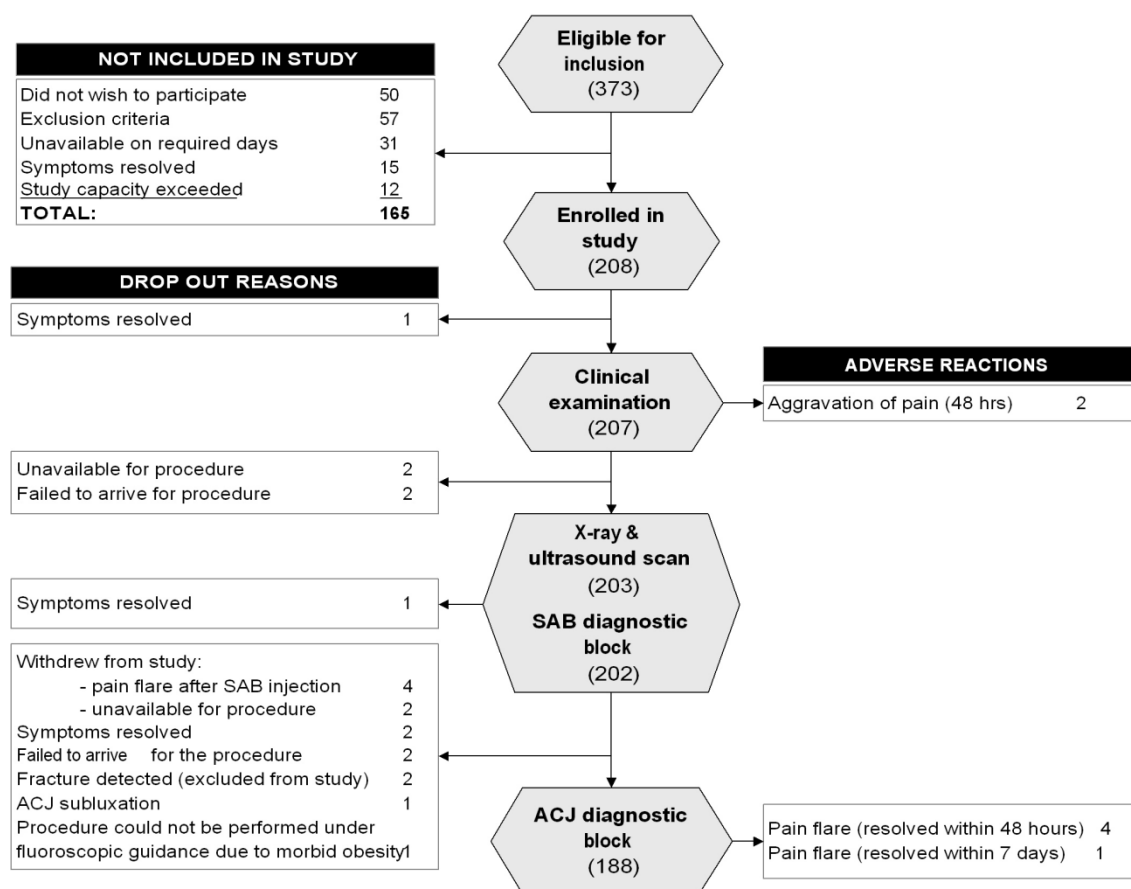
Figure 6.1. Flow chart of study procedures. Figure showing progression of participants through the study, dropout explanations and adverse reactions. SAB, subacromial bursa; ACJ, acromioclavicular joint.

Acromioclavicular Joint Diagnostic Block

One week following the imaging investigations and SAB injection, participants received a fluoroscopically guided injection of local anaesthetic into the ACJ. Participants were positioned supine with the arm in external rotation. Under aseptic conditions, a 22-gauge needle was inserted into the ACJ using a direct anterior approach. Iodinated contrast (0.5ml of Omnipaque 300 GE Healthcare) was introduced and fluoroscopic images used to confirm needle placement within the ACJ. Approximately 2mL of 1% lidocaine hydrochloride (Xylocaine[TM]) was then injected into the joint. The radiologist recorded whether the ACJ was successfully infiltrated and whether the injectate was contained within the joint.

Immediately prior to the injection, all participants were examined using up to six tests identified during the clinical examination as being provocative of typical symptoms. Pre-injection pain intensity was recorded for each clinical test on a 100mm

visual analogue scale (VAS; 0mm "no pain" and 100mm "worst imaginable pain"). Tests were repeated between 5 and 15 minutes following the injection and pain intensity scores recorded again. The percentage change in pain intensity was calculated for each index test and the average change in pain intensity from all clinical tests was calculated. Positive integers (+) indicate increased post-injection pain intensity, and negative integers (-) indicate decreased post-injection pain intensity. A positive anaesthetic response (PAR) was determined by 80% or more reduction in pain intensity post-injection. This is similar to the criteria for PAR used in other studies involving diagnostic blocks (Strobel et al., 2003) and was selected to maximise the level of certainty that the target structures were responsible for the majority of the symptoms.

The investigator performing the clinical examination and pre- and post-injection clinical tests (AC) was blinded to any diagnostic or treatment information from referring practitioners and to imaging results. Radiologists were not provided with any clinical information prior to the imaging or injection procedures.

Statistical Analysis

Due to the known limitations of VAS scales for measuring change in pain intensity when pre-injection pain levels are low (<20mm) (Bogduk, 2004b), only cases where pre-injection pain intensity exceeded 20mm were included in the analysis of anaesthetic response to diagnostic injections. The Fisher exact test was performed to identify variables associated with an 80% PAR to ACJ diagnostic block ($p \leq .200$), and odds ratios (OR) calculated for an 80% PAR using Statistical Package for the Social Sciences (SPSS version 17.0, IBM® Corporation 2010). The diagnostic accuracy of individual, and combinations of these clinical examination variables was assessed by calculating sensitivity, specificity, predictive values, positive likelihood ratios (+LR), negative likelihood ratios (-LR) and area under the receiver operator curve (ROC) with 95% confidence intervals (CI) using Confidence Interval Analysis software (Bryant, 2000).

Results

Three hundred and seventy three patients were referred to the study between July 2009 and June 2010 resulting in 208 participants being included in the study. Reasons for exclusion of patients in the study are presented in Figure 6.1. There were no significant differences between those included and excluded from the study with respect to age or gender. Those excluded from the study reported shorter duration of symptoms

(median 2 weeks; IQ range 4 weeks) (Mann-Whitney $p<0.001$). Demographic data for those who underwent the ACJ diagnostic block are presented in Table 6.1. Mean time between the clinical examination and ACJ diagnostic block was 11 days ($\pm$ 3 days), range 8 to 19 days.

Table 6.1. *Demographic Information*

Demographic information	ACJ completed (N=188) Mean (SD)	Range	PAR Group (n=22) Mean (SD)	NAR Group (n=131) Mean (SD)
Age (years)	42 (13)	18 - 81	41 (13)	43 (14)
Height (cm)	172 (10)	147 - 199	170 (11)	172 (10)
Weight (kg)	80.4 (16.7)	50.3 – 135.4	78.5 (16.1)	80.4 (17.0)
Symptom duration (weeks)*	7 (14)*	0 – 175	6 (18)*	8 (14)*
VAS (worst on 100mm scale)	62 (23)	3 – 100	59 (17)	65 (22)
VAS (average on 100mm scale)	36 (21)	1 - 100	33 (15)	39 (22)
SF8 physical component score (%)	44 (8)	23 – 61	45 (7)	44 (8)
SF8 mental component score (%)*	54 (11)*	27 – 66	57 (11)*	54 (10)*
SPADI pain score (%)	50 (22)	0 – 100	49 (15)	51 (21)
SPADI disability score (%)*	26 (30)*	0 – 96	26 (21)*	28 (30)*
SPADI total (%)	37 (20)	0 – 98	35 (13)	38 (21)
FABQ physical activity score (%)	65 (22)	0 – 100	64 (20)	65 (67)
FABQ work score (%)[a]*	21 (44)*	0 – 81	32 (47)*	21 (44)*
FABQ total score (%)[a]	41 (19)	0 – 87	45 (18)	41 (18)
% male gender	52		55	56
% right hand dominant	87		86	87
% dominant arm affected	52		36	55
% ACC claim	93		91	92
% physiotherapist referrals	97		100	96
% in paid employment	80		91	81
% on modified duties	10		5	12
% off work	3		9	3
% co-existent medical conditions	33		23	36
% smoker	20		9	22

Abbreviations. PAR, positive anaesthetic response ($\geq$80% post-injection reduction in pain intensity); NAR, negative anaesthetic response (<80% reduction in post-injection pain intensity); VAS, 100mm visual analogue pain score in previous 48 hours; SPADI, Shoulder Pain & Disability Index; FABQ, Fear Avoidance Beliefs Questionnaire; ACC, Accident Compensation Corporation

[a]only cases 'in paid employment' used in analysis
*median (interquartile range) values are presented. Variables were not normally distributed

Acromioclavicular Joint Diagnostic Block and Anaesthetic Response

One hundred and eighty eight participants received the ACJ diagnostic block. Drop-out explanations are presented in Figure 6.1. There were no differences in demographic or self-report questionnaire results between those who completed the study and those who dropped out ($p>0.05$). Average ACJ injection volume was 2.1mL (SD 0.7mL), and the injectate was contained within the ACJ in 174 cases (93%). Cases in which injectate was not contained within the ACJ (14), pre-injection pain intensity was less than 20mm on the VAS scale (21) and post-injection pain intensity was not

recorded (2) were excluded from the analysis. Post-injection reduction in average pain intensity was reported by 134 participants (88%) (range -1% to -100%), and post-injection increase in average pain intensity was reported by 18 participants (12%) (range +1% to +55%). One participant reported no change in post-injection pain intensity. A PAR (≥80% reduction in post-injection pain intensity) was reported by 22 of the 153 participants (14%). The distribution of diagnostic imaging results in the PAR and NAR groups, and in the group who reported a post-injection increase in pain is presented in Table 6.2.

Table 6.2. *Distribution of Diagnostic Imaging Results*

Diagnostic test results	Total identified	% in PAR group with test result	% in NAR group with test result	
			decreased pain intensity group	increased pain intensity group
	(N)	(n=22)	(n=113)	(n=18)
X-ray				
ACJ pathology	21	23	13	22
ACJ arthropathy	18	18	10	22
ACJ osteolysis	6	5	4	6
GHJ pathology	7	0	5	6
Rotator cuff calcification	19	5	9	39*
Ultrasound				
SAB pathology	105	55	71	56
Rotator cuff tear	46	14	31	28
Rotator cuff tendinosis	21	9	15	6
Rotator cuff calcification	35	18	20	44*
LHB tear or tendinosis	6	5	4	6
Biceps tendon sheath effusion	21	9	14	6
ACJ pathology	35	41	23	28
GHJ effusion	6	0	3	6

Abbreviations. ACJ, acromioclavicular joint; PAR, positive anaesthetic response (≥80% post-injection reduction in pain intensity); NAR, negative anaesthetic response (<80% reduction in post-injection pain intensity); GHJ, glenohumeral joint; SAB, subacromial bursa; LHB, long head of biceps
Note. Pathology subgroup totals may exceed composite pathology totals due to some cases identified in which multiple pathology was present.
*Significant difference in proportion of participants with and without diagnostic imaging finding ($p<0.05$).

Diagnostic Accuracy of the Clinical Examination

Nine clinical examination variables were associated with an 80% PAR ($p≤0.200$) (Table 6.3). A thickened or swollen ACJ had the highest OR for an 80% PAR (4.9; 95% CI 1.7, 14.4). For individual clinical variables, highest sensitivity was observed for pain referring below the elbow (1.00; 95% CI 0.84, 1.00) and highest specificity was

observed for repetitive mechanism of injury (0.90, 95% CI 0.84, 0.94). The highest +LR for individual clinical variables was recorded when passive external rotation performed at 90° abduction did not reproduce typical symptoms (2.83; 95% CI 1.56, 4.76) and the lowest -LR for an 80% PAR occurred when pain did not refer below the elbow (0.00; 95% CI 0.00, 0.92). For combinations of clinical examination findings, highest sensitivity (1.00; 95% CI 0.85, 1.00) and lowest -LR (0.00; 95% CI 0.00, 6.76) were observed when at least one of the nine tests was positive, and highest specificity (1.00; 95% CI 0.97, 1.00) and +LR (infinity; 95% CI 0.76, 428.00) occurred when eight of the nine tests were positive (Table 6.4). Area under the receiver operator curve for the total number of positive clinical tests was 0.727 ($p<0.05$; 95% CI: 0.585, 0.867) with the optimal diagnostic point identified when four positive clinical findings were present (sensitivity 0.77, specificity 0.69).

Discussion

The ability to identify patients likely to report a positive response to injection of local anaesthetic into the ACJ provides valuable diagnostic information for the primary care practitioner regarding the contribution of the ACJ to the patient's symptoms, and may provide clinical rationale for the use of targeted pain relief interventions including corticosteroid injections. It may also inform decision making regarding referral for more expensive or invasive investigation or treatment procedures. Combinations of nine history and physical examination features of the clinical examination were identified that were of diagnostic value for an 80% PAR to ACJ diagnostic block.

There were some limitations to consider in the current study. The false-positive rate for anaesthetic responses in peripheral joints including the ACJ has not been reported and results should be interpreted accordingly. Interobserver reliability of the history variables and observation of ACJ swelling or thickening also requires further evaluation. The use of strict cut-off criteria for a PAR may eliminate cases where the result may still produce a clinically meaningful outcome and ongoing analyses will be conducted in which various anaesthetic response levels will be used as outcome variables.

Table 6.3. *Diagnostic Accuracy of Individual Clinical Examination Variables for a Positive Response to Acromioclavicular Joint Diagnostic Block*

Clinical examination variables	Cell Counts				Diagnostic Accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)
Onset: repetitive activity	6	16	13	118	0.27 (0.13, 0.48)	0.90 (0.84, 0.94)	0.32 (0.15, 0.54)	0.88 (0.82, 0.93)	2.75 (1.15, 6.07)	0.81 (0.57, 0.98)	3.4* (1.1, 10.2)
No pain referred below elbow	20	0	105	23	1.00 (0.84, 1.00)	0.18 (0.12, 0.26)	0.16 (0.11, 0.23)	1.00 (0.86, 1.00)	1.22 (1.18, 1.34)	0.00 (0.00, 0.92)	0.84* (0.78, 0.91)
Pain does not disturb sleep	13	8	56	74	0.62 (0.41, 0.79)	0.57 (0.48, 0.65)	0.19 (0.11, 0.30)	0.90 (0.82, 0.95)	1.44 (0.91, 2.02)	0.67 (0.36, 1.07)	2.2 (0.8, 5.5)
ACJ thickened or swollen	15	5	47	77	0.75 (0.53, 0.89)	0.62 (0.53, 0.70)	0.24 (0.15, 0.36)	0.94 (0.87, 0.97)	1.98 (1.33, 2.70)	0.40 (0.28, 0.77)	4.9** (1.7, 14.4)
HBB – no pain	8	13	30	96	0.38 (0.21, 0.59)	0.76 (0.68, 0.83)	0.21 (0.11, 0.36)	0.88 (0.81, 0.93)	1.60 (0.82, 2.81)	0.81 (0.53, 1.07)	1.97 (0.75, 5.20)
No painful arc abduction	12	9	39	71	0.57 (0.37, 0.76)	0.65 (0.55, 0.73)	0.24 (0.14, 0.37)	0.89 (0.80, 0.94)	1.61 (0.98, 2.41)	0.66 (0.37, 1.02)	2.4 (0.9, 6.3)
PROM GHJ abduction – no pain	8	14	18	108	0.36 (0.20, 0.57)	0.86 (0.79, 0.91)	0.31 (0.17, 0.50)	0.89 (0.82, 0.93)	2.55 (1.23, 4.86)	0.74 (0.50, 0.95)	3.4* (1.3, 9.3)
PROM ER90^{0} – no pain	11	11	23	107	0.50 (0.31, 0.69)	0.82 (0.75, 0.88)	0.32 (0.19, 0.49)	0.91 (0.84, 0.95)	2.83 (1.56, 4.76)	0.61 (0.37, 0.85)	4.7** (1.8, 12.0)
PROM CB adduction – no pain	11	11	41	81	0.50 (0.31, 0.69)	0.66 (0.58, 0.74)	0.21 (0.12, 0.34)	0.88 (0.80, 0.93)	1.49 (0.87, 2.29)	0.75 (0.46, 1.08)	2.0 (0.8, 4.9)

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; ACJ, acromioclavicular joint; HBB, hand-behind-back; PROM, passive range of motion; GHJ, glenohumeral joint; ER90^{o}, external rotation performed in 90° of abduction; CB, cross-body

Note. Cell counts do not total 153 in some cases due to missing data

*p<0.05; ** p<0.01; *** p≤0.001.

Table 6.4. *Diagnostic accuracy of Combinations of Clinical Examination Variables for a Positive Response to Acromioclavicular Joint Diagnostic Block*

Number of positive clinical tests[a]	Cell counts				Diagnostic accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)
One or more	22	0	122	3	1.00 (0.85, 1.00)	0.02 (0.01, 0.07)	0.15 (0.10, 0.22)	1.00 (0.44, 1.00)	1.03 (1.02, 1.07)	0.00 (0.00, 6.76)	0.85 (0.79, 0.91)
Two or more	21	1	96	21	0.96 (0.78, 0.99)	0.18 (0.12, 0.26)	0.18 (0.12, 0.26)	0.96 (0.78, 0.99)	1.16 (0.95, 1.31)	0.25 (0.04, 1.28)	4.59 (0.59, 36.08)
Three or more	19	2	71	45	0.91 (0.71, 0.97)	0.39 (0.30, 0.48)	0.21 (0.14, 0.31)	0.96 (0.86, 0.99)	1.48 (1.13, 1.78)	0.25 (0.07, 0.77)	6.02* (1.34, 27.10)
Four or more	17	4	42	83	0.81 (0.60, 0.92)	0.66 (0.58, 0.74)	0.29 (0.19, 0.41)	0.95 (0.89, 0.98)	2.41 (1.67, 3.27)	0.29 (0.12, 0.61)	8.40*** (2.66, 26.54)
Five or more	15	7	23	101	0.68 (0.47, 0.84)	0.82 (0.74, 0.87)	0.40 (0.26, 0.55)	0.94 (0.87, 0.97)	3.68 (2.24, 5.76)	0.39 (0.20, 0.65)	9.41*** (3.45, 25.71)
Six or more	6	13	13	115	0.32 (0.15, 0.54)	0.90 (0.83, 0.94)	0.32 (0.15, 0.54)	0.90 (0.83, 0.94)	3.11 (1.31, 6.73)	0.76 (0.51, 0.95)	4.08* (1.33, 12.57)
Seven or more	3	18	5	125	0.14 (0.05, 0.35)	0.96 (0.91, 0.98)	0.38 (0.14, 0.69)	0.87 (0.81, 0.92)	3.71 (1.01, 12.77)	0.89 (0.68, 1.00)	4.17 (0.92, 18.94)
Eight or more	1	20	0	131	0.05 (0.01, 0.23)	1.00 (0.97, 1.00)	1.00 (0.21, 1.00)	0.87 (0.80, 0.91)	~ (0.76, 428.00)[†]	0.95 (0.83, 1.05) [†]	7.55 (5.02, 11.36)

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio

Note: Cell counts do not total 153 in some cases due to missing data

No participants had nine positive clinical tests

[a]Clinical tests: onset of pain due to repetitive activity; no pain referral below elbow; pain does not disturb sleep; ACJ thickened or swollen; no arc of pain during abduction; no symptom reproduction with hand-behind-back, PROM GHJ abduction, PROM external rotation (90° abduction), PROM cross-body adduction (in external rotation)

*p<0.05; ** p<0.01; *** p≤0.001.

[†]0.5 added to cells for estimate of confidence interval.

In agreement with previous reports in which external rotation of the shoulder was reported to help differentiate ACJ pain from intra-articular glenohumeral joint pathology (Cyriax, 1978), the absence of symptom reproduction during external rotation (at 90° abduction) in this study demonstrated high levels of specificity for an 80% PAR following ACJ diagnostic block. Reproduction of symptoms during the cross-body adduction test is reported to be diagnostic for pain arising from the ACJ (Cyriax, 1978), however in contrast, the absence of symptoms with the cross-body adduction test was associated with an 80% PAR in the current study. Cases were excluded from the analysis in whom injectate was observed to extravasate outside the ACJ capsule, as the capsuloligamentous disruption that commonly occurs following trauma to the ACJ (Rockwood, Williams, & Young, 1998) may have been present in these cases. It is possible that the direction of stress applied to the ACJ during the cross-body adduction test results in increased stress applied to ligamentous structures compared with structures involved in other ACJ pathologies.

Our results provide the first known reports of diagnostic accuracy of patient history variables for an 80% PAR to ACJ diagnostic block. A repetitive mechanism of injury demonstrated the highest specificity (90%) of all variables for an 80% PAR. A repetitive mechanism of injury was defined as the onset of pain during, or within 48 hours following a repetitive activity, within which time-frame no other specific cause was identified. Repetitive microtrauma is also a well documented cause of osteolysis of the distal clavicle that results in pain in the region of the ACJ (Cahill, 1992; Kaplan & Resnick, 1986). The absence of pain radiating below the level of the elbow demonstrated the highest sensitivity of any individual variable (100%) meaning an 80% PAR could effectively be ruled out if a participant reported pain extending below the elbow. Previous reports also indicate that pain referral extending below the elbow following irritation of the ACJ is rare (Gerber et al., 1998).

When compared with individual clinical examination variables, combinations of the nine clinical variables demonstrated improved the specificity and increased likelihood of an 80% PAR. Highest levels of specificity (1.00) and +LR (infinity) were observed when eight clinical findings were present, however only one participant satisfied this criterion limiting its clinical application to a small proportion of patients. When five or more clinical findings were present specificity still exceeded 80% with a +LR of 3.68 and a lower confidence limit that exceeded 2.00 representing confidence that someone with at least five findings would report an 80% PAR.

The ability to rule-out an 80% PAR to ACJ diagnostic block enables identification of patients in whom the ACJ may not be responsible for the majority of symptoms, and may identify patients unlikely to report high levels of pain relief following targeted pain relief interventions such as corticosteroid injections. If a participant reported less than three of the nine clinical findings, there was at least a 91% probability that an 80% PAR could be ruled out. Although sensitivity improved to 100% and the -LR reduced to 0.00 when only one clinical finding was present, the upper 95% confidence limit for the -LR was 6.76, reducing confidence in this clinical decision. Four or more clinical findings represented the 'optimal diagnostic point' for both sensitivity and specificity, however the lower confidence limits for sensitivity and specificity were 60% and 58% respectively rendering these diagnostic characteristics equivocal in this situation. In clinical practice patients with four positive clinical findings may benefit from an intra-articular ACJ diagnostic injection to determine the anaesthetic response.

Almost the same proportion of participants who reported 80% pain relief following the ACJ diagnostic block reported provocation of pain during the ACJ diagnostic block procedure, with elevated pain intensity levels persisting into the post-injection clinical reassessment period. This was an unexpected finding, and it is possible that short-term provocation of pain from needle entry into a pathologic synovial joint capsule may occur and mask any medium-long term pain relief obtained from the anaesthetic. Such provocation in response to the injection has been reported in the literature in relation to spinal pain (Schwarzer et al., 1995), and although this has been deemed an unreliable indicator of lumbar zygapophyseal joint pathology (Schwarzer, Derby, et al., 1994), it is possible that symptom provocation during ACJ injection may be associated with pathologic lesions affecting this joint. More advanced imaging investigations such as magnetic resonance imaging would be required to evaluate this theory.

Conclusion

In conclusion, this study provides preliminary evidence of the importance of aspects of patient history in predicting an 80% PAR to ACJ diagnostic block in patients with shoulder pain recruited from a primary care setting. Compared with individual clinical examination features, combinations of nine clinical examination variables improved the ability to predict an 80% PAR to ACJ diagnostic block when at least five clinical features were present. Such findings may aid clinical decision making regarding the use of targeted pain relief interventions, and referral for more expensive or invasive treatment procedures.

6.2. Added Value of Imaging Findings for Predicting a Positive Response to Guided Acromioclavicular Joint Diagnostic Block

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Taylor, S. The added value of imaging findings for predicting a positive response to guided acromioclavicular joint diagnostic block.

The results presented in the following manuscript have been formatted in manuscript style and will be submitted for publication pending further investigation.

Abstract

Background: Clinical predictors of a PAR to ACJ diagnostic block were identified in the previous section. In clinical practice imaging is often used to assist in the diagnosis of subacromial disorders, however it is unknown whether imaging findings improve the ability to identify those likely to report a PAR, indicating the ACJ as the pain source.

Objectives: To assess the relationship between specific ultrasound imaging findings of ACJ pathology and the response to guided ACJ diagnostic block, and evaluate the added diagnostic value of these imaging findings when combined with clinical examination findings, for predicting a positive anaesthetic response (PAR).

Methods: Consecutive patients with shoulder pain underwent a standardised clinical examination, shoulder x-ray series and diagnostic ultrasound. Results were compared with the response to a diagnostic block of xylocaine^{TM} injected into the ACJ under fluoroscopic guidance. The diagnostic accuracy of clinical examination variables for a PAR (≥80% reduction in post-injection pain intensity) was compared with the diagnostic accuracy of combinations of clinical and imaging variables associated with a PAR (p<0.10).

Results: A PAR was reported by 14% of participants. The strongest ultrasound imaging predictors of a PAR were ACJ capsular hypertrophy, any ACJ pathology and an intact rotator cuff (no tear identified). Highest sensitivity for a PAR was observed for the clinical examination finding of no referral of pain below the elbow (100%; 95% CI 0.84, 1.00). The presence of ACJ capsular hypertrophy on ultrasound was the strongest predictor of a PAR (OR 17.57; 95% CI 4.69, 65.87; p<0.001) and when combined with

clinically observed swelling or thickening of the ACJ resulted in 100% specificity (95% CI 0.97, 1.00) and post-test probability of 100% (95% CI 0.65, 1.00) for a PAR.

Conclusion: Preliminary results suggest that imaging findings do not improve the ability to rule-out a PAR compared with clinical examination findings alone, however ACJ capsular hypertrophy on ultrasound does improve the ability to rule-in an 80% PAR when combined with clinical examination findings. Verification of these results is required in future studies.

Introduction

Disorders of the acromioclavicular joint (ACJ) are a common cause of shoulder pain, affecting patients of all ages and levels of activity (Shaffer, 1999). Identification of the ACJ as the primary source of pain is important to enable efficient application of appropriate treatment interventions, as well as to inform decisions regarding referral for further medical or imaging investigations or specialist consultation.

The diagnosis of shoulder pain begins with a clinical examination, and imaging investigations such as x-ray, diagnostic ultrasound and magnetic resonance imaging (MRI) are also frequently used to identify a pathoanatomic lesion that may represent the source of pain. Previous studies demonstrate that there is a high prevalence of ACJ pathology on imaging in asymptomatic individuals with asymptomatic ACJ arthritic changes identified on magnetic resonance imaging (MRI) in 93% of individuals over the age of 30 years (Shubin Stein et al., 2001). The high prevalence of asymptomatic ACJ pathology complicates interpretation of imaging findings with respect to the identification of symptomatic ACJ conditions.

However, symptomatic ACJ lesions have been associated with higher grades of ACJ degenerative changes on MRI (Shubin Stein et al., 2006). Capsular hypertrophy, joint effusion as well as more advanced changes such as osteophytes, subchondral cysts, and reactive marrow oedema are also reported to be associated with ACJ pain diagnosed by a positive response to injection of local anaesthetic (Shubin Stein et al., 2006; Strobel et al., 2003). However, MRI is not widely available to primary care practitioners in many countries, and is an expensive investigation. The ability to diagnose painful ACJ conditions using combinations of clinical examination and imaging modalities that are less expensive, such as diagnostic ultrasound, would improve accessibility to these investigations and reduce costs of, and necessity for investigations such as MRI.

Diagnostic ultrasound scans are being increasingly used to aid in the diagnosis of shoulder pain (Awerbuch, 2008). While more commonly used for imaging of subacromial pathologies such as subacromial bursa and rotator cuff pathology, ultrasound is also able to visualise ACJ capsular structures and bone contour profiles to identify hypertrophic capsular tissue or bony cortical erosions that may indicate arthropathic joint disease, as well as enable dynamic assessment of the ACJ for suspected joint instability (Martinoli et al., 2003). However, the relevance of such ultrasound findings to the symptoms of ACJ pain diagnosed using the accepted reference standard (response to injection of local anaesthetic) has not been extensively investigated.

The aim of this study was to evaluate whether any specific ultrasound findings were associated with a positive response to guided ACJ diagnostic block, and to evaluate the added diagnostic value of any findings when combined with clinical examination findings, for predicting a PAR to ACJ diagnostic block.

Methods

Design, Sampling and Recruitment

Design, sampling and recruitment were the same as reported in the previous section (p160).

Diagnostic Imaging and Acromioclavicular Joint Diagnostic Block

Clinical examination procedures, x-ray, diagnostic ultrasound scan and the fluoroscopic guided ACJ diagnostic block procedures were described in Chapter 4 (p104). Information regarding the presence of specific ACJ pathology on ultrasound was either prospectively recorded on data collection forms, or was retrieved retrospectively from sonographer worksheets used as part of normal practice procedures. All other x-ray and ultrasound variables were recorded on a standardised data collection form provided to the sonographer and radiologist. The examiner who performed pre- and post-injection clinical examination testing was blinded to imaging results and the radiologist performing the ACJ diagnostic block was blinded to the clinical examination findings.

Statistical Methods

The clinical examination variables identified in the previous section that were associated with a PAR to ACJ diagnostic block ($p<0.05$) were selected for inclusion in diagnostic test combinations (Table 6.3). Fisher's test was used to assess the association between x-ray and ultrasound variables and a PAR to ACJ diagnostic block. Imaging variables associated with a PAR ($p<0.10$) were selected for inclusion in the diagnostic accuracy analysis of variable combinations. A less stringent Fisher's test p-value ($p\leq0.10$) was selected to avoid elimination of an imaging variable with stricter cut-off criteria which can be affected by low prevalence of the condition.

The diagnostic accuracy of individual and combinations of clinical examination variables only (based upon minimum numbers of positive tests) was calculated. Cases were then identified in which both the clinical examination and imaging findings were present and diagnostic accuracy for a PAR was calculated. To evaluate the added diagnostic value of the imaging findings, diagnostic accuracy results of the clinical examination test criteria alone were compared with the accuracy of combined clinical examination and imaging criteria.

Results

The number of participants included in the study and descriptive information are presented in Figure 6.1 (p162). Demographic information was presented in Table 6.1 (p164).

One hundred and eighty eight participants received the ACJ diagnostic block. Drop-out explanations are presented in Figure 6.1. Average ACJ injection volume was 2.1mL (SD 0.7mL), and the injectate was contained within the ACJ in 174 cases (93%). A PAR ($\geq80\%$ reduction in post-injection pain intensity) was reported by 22 of the 153 participants (14%). The distribution of diagnostic imaging results in the PAR and NAR groups, and in the group who reported a post-injection increase in pain is presented in Table 6.2 (p165).

Clinical Examination and Imaging Variables

Repetitive onset of activity, no referral of pain below the elbow, thickened or swollen ACJ and the absence of symptom reproduction during passive ROM glenohumeral abduction or external rotation (at 90° abduction) were associated with a PAR to ACJ diagnostic block ($p<0.05$) and were included in the analysis.

The relationships between x-ray and ultrasound scan variables and a PAR to ACJ diagnostic block are presented in Table 6.5. Capsular hypertrophy of the ACJ reported on ultrasound was the strongest independent predictor of an 80% PAR (OR 17.57; 95% CI 4.69, 65.87; p=0.000). Any ACJ pathology and the absence of a rotator cuff tear (any rotator cuff component) on ultrasound were also associated with a PAR to ACJ diagnostic block (p<0.10). Diagnostic accuracy of individual imaging findings are presented in Table 6.6. Diagnostic accuracy of individual (Table 6.7), and combinations of clinical examination variables and imaging findings (Table 6.8) are presented below.

The absence of referred pain below the elbow demonstrated highest sensitivity (1.00) and repetitive mechanism of pain onset demonstrated highest specificity (0.90). For clinical test combinations, highest sensitivity was observed when at least one of the five clinical features was not present (0.94) and highest specificity when three or more of the five features were present (0.92). For individual imaging variables, highest sensitivity was observed for 'no rotator cuff tear on ultrasound' (0.86; -LR 0.42) and highest specificity (0.97; +LR 11.55) was observed for ACJ capsular hypertrophy reported on ultrasound. No cases were identified in which four or five clinical features were present.

Added Diagnostic Value of Imaging Findings

When imaging variables were added to individual clinical examination variables, sensitivity ranged from 0.00 (no pain referral below elbow with ACJ pathology and no rotator cuff pathology) to 0.70 (thickened or swollen ACJ with no rotator cuff tear on ultrasound) and specificity ranged from 0.73 (thickened or swollen ACJ with no rotator cuff tear on ultrasound) to 1.00 (ACJ capsular hypertrophy on ultrasound combined with either repetitive onset, thickened or swollen ACJ or absence of symptom provocation with GHJ abduction).

Table 6.5. *Distribution of Imaged Pathology in ACJ Diagnostic Block PAR and NAR Groups.*

Diagnostic test results	Total identified (N)	% in PAR group with pathology (n=22)	% in NAR group with pathology (n=131)	OR (95% CI)
X-ray				
ACJ pathology	22	23	12	2.11 (0.69, 6.52)
ACJ arthropathy	18	18	11	1.86 (0.55, 6.27)
ACJ osteolysis	6	5	4	1.20 (0.13, 10.79)
os acromiale	4	0	3	1.17 (1.10, 1.25)
GHJ pathology	7	0	5	1.18 (1.10, 1.26)
Rotator cuff calcification	19	5	14	0.30 (0.04, 2.36)
supraspinatus	11	5	8	0.58 (0.07, 4.74)
infraspinatus	7	0	5	1.18 (1.10, 1.26)
subscapularis	6	0	5	1.18 (1.10, 1.26)
Ultrasound				
ACJ pathology	35	41	20	2.72 (1.05, 7.04)*
capsular hypertrophy	12	36	3	17.57 (4.69, 65.87)***
bony irregularity	7	5	5	0.97 (0.11, 8.46)
joint space widening	3	0	2	1.18 (1.10, 1.26)
SAB pathology[a]	51	23	35	0.54 (0.19, 1.57)
fluid/effusion	15	0	12	1.19 (1.11, 1.28)
calcification	3	0	2	1.17 (1.10, 1.25)
thickened ($\geq$2mm)	39	23	26	0.84 (0.29, 2.45)
Bursal bunching				
acromion	63	36	45	0.71 (0.28, 1.81)
CAL[b]	38	56	51	1.21 (0.30, 4.92)
Rotator cuff tear (any)	46	14	33	0.32 (0.09, 1.15)*
Supraspinatus pathology	66	36	44	0.72 (0.28, 1.83)
calcification	22	18	14	1.40 (0.42, 4.60)
tendinosis	20	9	14	0.63 (0.14, 2.92)
tear	38	14	27	0.43 (0.12, 1.55)
Infraspinatus pathology	12	0	9	1.19 (1.10, 1.27)
calcification	9	0	7	1.18 (1.10, 1.27)
tendinosis	1	0	1	1.17 (1.10, 1.25)
tear	3	0	2	1.17 (1.10, 1.25)
Subscapularis pathology	22	5	16	0.25 (0.03, 1.96)
calcification	15	5	11	0.40 (0.05, 3.19)
tendinosis	1	0	1	1.17 (1.10, 1.25)
tear	10	0	8	1.18 (1.10, 1.27)
LHB tear or tendinosis	6	5	4	1.25 (0.14, 11.26)
LHB sheath effusion	21	9	15	0.59 (0.13, 2.73)
GHJ effusion	6	0	5	1.18 (1.10, 1.26)

Abbreviations. PAR, positive anaesthetic response; NAR, negative anaesthetic response; ACJ, acromioclavicular joint; GHJ, glenohumeral joint; SAB, subacromial bursa; CAL, coracoacromial ligament; PTT, partial thickness tear; FTT, full thickness tear; LHB, long head of biceps tendon.
[a] SAB pathology included: thickening $\geq$2mm, calcification, bursal fluid or effusion.
[b] bunching under the CAL only assessed in 93 cases. (PAR n=26; NAR n=67)
*$p<0.05$; ** $p<0.01$; *** $p\leq0.001$.

Table 6.6. *Diagnostic Accuracy of Ultrasound Imaging Variables for Positive Response to Acromioclavicular Joint Diagnostic Block*

Clinical examination variables	Cell Counts				Diagnostic Accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)
US: ACJ pathology	9	13	26	102	0.41 (0.23, 0.61)	0.80 (0.72, 0.86)	0.26 (0.14, 0.42)	0.89 (0.82, 0.93)	2.01 (1.05, 3.51)	0.74 (0.48, 0.98)	2.72 (1.05, 7.04)
US: ACJ capsular hypertrophy	8	14	4	123	0.36 (0.20, 0.57)	0.97 (0.92, 0.99)	0.67 (0.39, 0.86)	0.90 (0.84, 0.94)	11.55 (3.94, 33.20)	0.66 (0.44, 0.83)	17.57*** (4.69, 65.87)
US: No rotator cuff tear	19	3	88	43	0.86 (0.67, 0.95)	0.33 (0.25, 0.41)	0.18 (0.12, 0.26)	0.94 (0.83, 0.98)	1.29 (0.98, 1.53)	0.42 (0.14, 1.06)	3.10 (0.87, 11.03)

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; US, ultrasound; ACJ, acromioclavicular joint.

Note. Cell counts do not total 153 in some cases due to missing data

*p<0.05; ** p<0.01; *** p≤0.001.

Table 6.7. *Diagnostic Accuracy of Individual Clinical Examination and Imaging Variables for Positive Response to Acromioclavicular Joint Diagnostic Block*

Clinical examination variables	Cell Counts				Diagnostic Accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)
Onset: repetitive activity	6	16	13	118	0.27 (0.13, 0.48)	0.90 (0.84, 0.94)	0.32 (0.15, 0.54)	0.88 (0.82, 0.93)	2.75 (1.15, 6.07)	0.81 (0.57, 0.98)	3.4* (1.1, 10.2)
US: ACJ pathology	2	20	2	126	0.09 (0.03, 0.28)	0.98 (0.95, 1.00)	0.50 (0.15, 0.85)	0.86 (0.80, 0.91)	5.82 (1.05, 31.22)	0.92 (0.73, 1.00)	6.3 (0.84, 47.30)
US: ACJ capsular hypertrophy	2	20	0	127	0.09 (0.03, 0.28)	1.00 (0.97, 1.00)	1.00 (0.34, 1.00)	0.86 (0.80, 0.91)	~ (3.04, ~)	0.91 (0.93, ~)	7.4* (4.9, 11.1)
US: No rotator cuff tear	5	17	12	119	0.23 (0.10, 0.43)	0.91 (0.85, 0.95)	0.29 (0.13, 0.53)	0.88 (0.81, 0.92)	2.48 (0.96, 5.89)	0.85 (0.62, 1.01)	2.9 (0.9, 9.3)
No pain referred below elbow	20	0	105	23	1.00 (0.84, 1.00)	0.18 (0.12, 0.26)	0.16 (0.11, 0.23)	1.00 (0.86, 1.00)	1.22 (1.18, 1.34)	0.00 (0.00, 0.92)	0.84* (0.78, 0.91)
US: ACJ pathology	0	20	5	121	0.00 (0.00, 0.16)	0.96 (0.91, 0.98)	0.00 (0.00, 0.43)	0.86 (0.79, 0.91)	0.00 (0.00, 4.36)	1.04 (1.04, 1.10)	1.2 (1.1, 1.3)
US: ACJ capsular hypertrophy	0	20	0	125	-	-	-	-	-	-	-
US: No rotator cuff tear	0	20	13	115	0.00 (0.00, 0.16)	0.90 (0.83, 0.94)	0.00 (0.00, 0.23)	0.85 (0.78, 0.90)	0.00 (0.00, 1.64)	1.11 (1.10, 1.20)	1.2 (1.1, 1.3)
ACJ thickened or swollen	15	5	47	77	0.75 (0.53, 0.89)	0.62 (0.53, 0.70)	0.24 (0.15, 0.36)	0.94 (0.87, 0.97)	1.98 (1.33, 2.70)	0.40 (0.28, 0.77)	4.9** (1.7, 14.4)
US: ACJ pathology	7	13	9	112	0.35 (0.18, 0.57)	0.93 (0.87, 0.96)	0.44 (0.23, 0.67)	0.90 (0.83, 0.94)	4.71 (1.96, 10.66)	0.70 (0.47, 0.89)	6.7** (2.2, 21.0)
US: ACJ capsular hypertrophy	7	13	0	120	0.35 (0.18, 0.57)	1.00 (0.97, 1.00)	1.00 (0.65, 1.00)	0.90 (0.84, 0.94)	~ (11.18, ~)	0.65 (0.68, ~)	10.2*** (6.1, 17.1)
US: No rotator cuff tear	14	6	33	91	0.70 (0.48, 0.86)	0.73 (0.65, 0.80)	0.30 (0.19, 0.44)	0.94 (0.87, 0.97)	2.63 (1.66, 3.86)	0.41 (0.20, 0.72)	6.4*** (2.3, 18.1)

	TP	FN	FP	TN	Sensitivity (CI)	Specificity (CI)	PPV (CI)	NPV (CI)	+LR (CI)	-LR (CI)	OR (CI)
PROM GHJ abd – no pain	8	14	18	108	0.36 (0.20, 0.57)	0.86 (0.79, 0.91)	0.31 (0.17, 0.50)	0.89 (0.82, 0.93)	2.55 (1.23, 4.86)	0.74 (0.50, 0.95)	3.4* (1.3, 9.3)
US: ACJ pathology	4	18	1	122	0.18 (0.07, 0.39)	0.99 (0.96, 1.00)	0.80 (0.38, 0.96)	0.87 (0.81, 0.92)	22.36 (3.46, 143.98)	0.83 (0.62, 0.94)	27.1** (2.9, 256.3)
US: ACJ capsular hypertrophy	4	18	0	122	0.18 (0.07, 0.39)	1.00 (0.97, 1.00)	1.00 (0.51, 1.00)	0.87 (0.81, 0.92)	~ (0.59, ~)	0.82 (0.85, ~)	7.8*** (5.1, 12.0)
US: No rotator cuff tear	8	14	14	112	0.36 (0.20, 0.57)	0.89 (0.82, 0.93)	0.36 (0.20, 0.57)	0.89 (0.82, 0.93)	3.27 (1.52, 6.56)	0.72 (0.48, 0.91)	4.6** (1.6, 12.8)
PROM ER90^0 – no pain	11	11	23	107	0.50 (0.31, 0.69)	0.82 (0.75, 0.88)	0.32 (0.19, 0.49)	0.91 (0.84, 0.95)	2.83 (1.56, 4.76)	0.61 (0.37, 0.85)	4.7** (1.8, 12.0)
US: ACJ pathology	3	19	3	124	0.14 (0.05, 0.33)	0.98 (0.93, 0.99)	0.50 (0.19, 0.81)	0.87 (0.80, 0.91)	5.77 (1.37, 23.32)	0.89 (0.68, 0.98)	6.5* (1.2, 34.7)
US: ACJ capsular hypertrophy	3	19	1	125	0.14 (0.05, 0.33)	0.99 (0.96, 1.00)	0.75 (0.30, 0.95)	0.87 (0.80, 0.91)	17.18 (2.52, 116.05)	0.87 (0.67, 0.96)	19.7** (2.0, 199.6)
US: No rotator cuff tear	11	11	16	114	0.50 (0.31, 0.69)	0.88 (0.81, 0.92)	0.41 (0.25, 0.59)	0.91 (0.85, 0.95)	4.06 (2.14, 7.33)	0.57 (0.34, 0.80)	7.1*** (2.7, 19.1)

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; US, ultrasound; ACJ, acromioclavicular joint; PROM, passive range of motion; GHJ abd, glenohumeral joint abduction; ER90^o, external rotation performed in 90° of abduction

Note. Cell counts do not total 153 in some cases due to missing data

- no cases identified, values could not be calculated

*p<0.05; ** *p*<0.01; *** *p*≤0.001.

Table 6.8. *Diagnostic Accuracy of Combinations of Clinical Examination and Imaging Variables for Positive Response to Acromioclavicular Joint Diagnostic Block*

Clinical examination variable combinations	Cell Counts				Diagnostic Accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)
1 or more tests	17	1	74	41	0.94 (0.74, 0.99)	0.36 (0.28, 0.45)	0.19 (0.12, 0.28)	0.98 (0.88, 1.00)	1.47 (1.13, 1.74)	0.16 (0.03, 0.74)	9.4* (1.2, 73.4)
US: ACJ pathology	6	12	14	99	0.33 (0.16, 0.56)	0.88 (0.80, 0.93)	0.30 (0.15, 0.52)	0.89 (0.82, 0.94)	2,69 (1.15, 5.69)	0.76 (0.50, 0.97)	3.5* (1.1, 10.9)
US: ACJ capsular hypertrophy	6	12	0	112	0.33 (0.16, 0.56)	1.00 (0.97, 1.00)	1.00 (0.61, 1.00)	0.90 (0.84, 0.94)	~ (9.94, ~)	0.67 (0.70, ~)	10.3*** (6.0, 17.7)
US: No rotator cuff tear	16	2	51	64	0.89 (0.67, 0.97)	0.56 (0.47, 0.64)	0.24 (0.15, 0.35)	0.97 (0.90, 0.99)	2.00 (1.45, 2.57)	0.20 (0.06, 0.60)	10.0*** (2.2, 45.7)
2 or more tests	10	8	31	84	0.56 (0.34, 0.75)	0.73 (0.64, 0.80)	0.24 (0.14, 0.39)	0.91 (0.84, 0.96)	2.06 (1.17, 3.26)	0.61 (0.33, 0.93)	3.4* (1.2, 9.4)
US: ACJ pathology	4	14	4	109	0.22 (0.09, 0.45)	0.97 (0.91, 0.99)	0.50 (0.22, 0.79)	0.89 (0.82, 0.93)	6.28 (1.80, 20.75)	0.81 (0.57, 0.95)	7.8* (1.7, 34.7)
US: ACJ capsular hypertrophy	4	14	0	112	0.22 (0.09, 0.45)	1.00 (0.97, 1.00)	1.00 (0.51, 1.00)	0.89 (0.82, 0.93)	~ (6.60, ~)	0.78 (~, ~)	9.0*** (5.5, 14.7)
US: No rotator cuff tear	10	8	23	92	0.56 (0.34, 0.75)	0.80 (0.72, 0.86)	0.30 (0.17, 0.47)	0.92 (0.85, 0.96)	2.78 (1.52, 4.62)	0.56 (0.31, 0.84)	5.0** (1.8, 14.1)
3 or more tests	4	14	9	106	0.22 (0.09, 0.45)	0.92 (0.86, 0.96)	0.31 (0.13, 0.58)	0.88 (0.81, 0.93)	2.84 (0.98, 7.52)	0.84 (0.59, 1.00)	3.4 (0.91, 12.4)
US: ACJ pathology	1	17	0	113	0.06 (0.01, 0.26)	1.00 (0.97, 1.00)	1.00 (0.21, 1.00)	0.87 (0.80, 0.92)	~ (1.65, ~)	0.94 (0.98, ~)	7.6 (4.9, 11.9)
US: ACJ capsular hypertrophy	1	17	0	112	0.06 (0.01, 0.26)	1.00 (0.97, 1.00)	1.00 (0.21, 1.00)	0.87 (0.80, 0.92)	~ (1.63, ~)	0.94 (0.98, ~)	7.6 (4.9, 11.8)
US: No rotator cuff tear	4	14	7	108	0.22 (0.09, 0.45)	0.94 (0.88, 0.97)	0.36 (0.15, 0.65)	0.89 (0.82, 0.93)	3.65 (1.20, 10.25)	0.83 (0.58, 0.98)	4.4* (1.1, 17.0)

	TP	FN	FP	TN							
4 or more tests	0	18	0	115	-	-	-	-	-	-	-
US: ACJ pathology	0	18	0	113	-	-	-	-	-	-	-
US: ACJ capsular hypertrophy					-	-	-	-	-	-	-
US: No rotator cuff tear	0	18	0	115	-	-	-	-	-	-	-
5 tests	0	18	0	115	-	-	-	-	-	-	-
US: ACJ pathology	0	18	0	113	-	-	-	-	-	-	-
US: ACJ capsular hypertrophy					-	-	-	-	-	-	-
US: No rotator cuff tear	0	18	0	115	-	-	-	-	-	-	-

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; US, ultrasound; ACJ, acromioclavicular joint

Note. Cell counts do not total 153 in some cases due to missing data.

- no cases identified, values could not be calculated

*p<0.05; ** p<0.01; *** p≤0.001.

When imaging findings were added to combinations of clinical examination findings, sensitivity ranged from 0.06 (3 or more positive tests and ACJ pathology or capsular hypertrophy also reported on ultrasound) to 0.89 (1 or more positive test and no rotator cuff tear on ultrasound). Specificity for clinical test combinations when imaging findings were added ranged from 0.56 (1 or more positive test and no rotator cuff tear on ultrasound) to 1.00 (at least one, two or three clinical examination findings and ACJ capsular hypertrophy also reported on ultrasound). A graphical presentation of the impact of imaging findings on the post-test probability for a PAR compared with clinical examination features alone is presented in Figure 6.2. The largest increase in post-test probability was observed when ACJ capsular hypertrophy was reported and only one or two clinical tests were positive (post-test probability 100%). Diagnostic accuracy statistics could not be calculated for four or five positive tests due to no cases being identified that fitted these criteria.

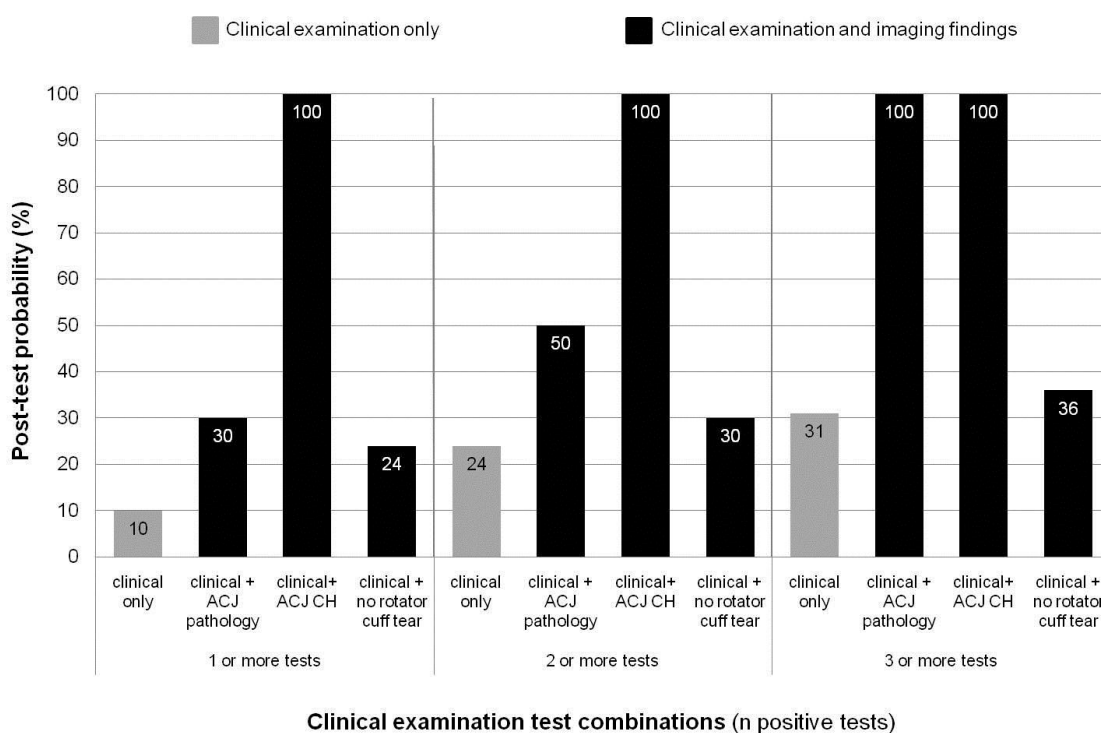*Figure 6.2.* Change in post-test probability for a PAR following ACJ diagnostic block with addition of imaging findings to clinical examination findings. Graph showing change in post-test probability (%) when ultrasound imaging findings are added to clinical test findings. ACJ, acromioclavicular joint; CH, capsular hypertrophy. No cases were identified in which four or five clinical examination tests were positive.

Discussion

This study aimed to investigate the added diagnostic value of imaging findings for identifying those participants likely to report a PAR to ACJ diagnostic block. Pathological ACJ changes that can be observed on ultrasound include widened joint space, clavicular displacement, soft tissue haematoma, cortical erosions, joint effusion, soft tissue swelling, cysts, capsular hypertrophy, calcification and intra-articular loose bodies (Martinoli et al., 2003). Pathological changes affecting the ACJ, particularly capsular hypertrophy, were shown to be associated with and to improve the ability to predict an 80% PAR to ACJ diagnostic block in this study.

Imaging Findings

The presence of ACJ capsular hypertrophy on ultrasound was the strongest independent predictor of a PAR (OR 17.57: +LR 11.55), and demonstrated higher specificity as an independent variable (0.97) than any of the individual or clinical test combinations. When combined with clinical examination findings, identification of ACJ capsular hypertrophy on ultrasound increased specificity of many individual and clinical test combinations to 100%, and increased the post-test probability from a maximum of 31% (three clinical features) to 100% when this imaging finding was added to the clinical features. While no other studies were identified comparing pathological ACJ findings on ultrasound with results of other reference standard procedures, ACJ capsular hypertrophy on MRI has been reported as one of the most sensitive (0.73) pathological ACJ variables for a PAR at the 70% pain relief standard (Strobel et al., 2003). In contrast, the specificity (0.51) and PPV (0.30) for the 70% PAR were lower than found in our study. This may be explained by the difference in population, and may be partially due to the different anaesthetic response criterion. Diagnostic ultrasound is also limited in its ability to visualise the inferior aspect of the ACJ due to overlying bone (acromion and humeral head). This may underestimate the prevalence of ACJ pathology compared with MRI or surgical findings, however results indicate this variable may represent a promising tool for identifying symptomatic ACJ pathology.

Added Diagnostic Value of Imaging Findings

When added to the clinical examination variables, ACJ capsular hypertrophy improved the probability of a PAR to 100% for several individual, and all combinations of clinical examination variables in which diagnostic accuracy could be calculated. The largest improvement in post-test probability was observed when ACJ capsular

hypertrophy reported on ultrasound was added to the clinical appearance of a thickened or swollen ACJ (62% to 100%) with the lower confidence limit for the post-test probability (65%) still representing a meaningful increase in the probability of a PAR compared with a pre-test probability (prevalence) of 14%. A large increase in post-test probability was also observed when any one or more of the clinical examination features was present (post-test probability 36%) and ACJ capsular hypertrophy was subsequently identified on ultrasound (post-test probability 100%).

Although ACJ capsular hypertrophy improved the post-test probability for a PAR, this was reported on ultrasound in only 8% of cases (12/153 participants). The low prevalence of this finding resulted in identification of only a small number of participants who satisfied clinical examination and imaging criteria that included this finding. As a result, only 1-5% of participants satisfied clinical examination and imaging criteria with high specificity for a PAR. In contrast, the prevalence of the composite imaging variable that included any ACJ pathology (capsular hypertrophy, bony irregularity, joint space widening or joint effusion) was higher, reported in 35 cases (23%). When combined with the absence of symptom provocation during passive GHJ abduction (post-test probability 30%), the report of 'any' ACJ pathology on ultrasound was also highly specific (0.99; 95% CI 0.96, 1.00) for a PAR, and resulted in a large increase in post-test probability to 80% (+LR 22.36). The finding of any ACJ pathology on ultrasound may therefore be applicable to a larger proportion of patients in the clinical setting. However, when ACJ capsular hypertrophy is specifically reported, it is likely to further improve the ability to predict an 80% PAR.

Although the combination of clinical examination and imaging findings were able to rule-in a PAR in some participants, our results suggest that clinical features alone were better able to rule-out an 80% PAR following ACJ diagnostic block. A PAR could be ruled-out with a moderate to high level of confidence when the participant reported referral of pain below the level of the elbow (sensitivity 1.00, -LR 0.00), or when at least one of the five clinical tests was not positive (sensitivity 0.94; -LR 0.16). The addition of imaging findings to these two variables resulted in lower sensitivity, most likely due to the low prevalence of participants who met the criteria that also required positive imaging findings. Highest sensitivity for combined clinical examination and imaging variables reached only 0.70 (thickened or swollen ACJ and no rotator cuff tear on ultrasound), however sensitivity for a PAR was higher with the clinical observation alone (thickened or swollen ACJ, sensitivity 0.75). These results suggest that clinical

examination tests alone may be a more effective screening tool for symptomatic ACJ pathology at the 80% pain relief standard than the combination of clinical examination and imaging findings.

Evidence of an intact rotator cuff on ultrasound, did not substantially alter the post-test probability of a PAR, or a NAR compared with clinical tests alone. While specificity of clinical tests for a PAR did improve slightly when an intact rotator cuff was reported on ultrasound, post-test probabilities for a PAR remained largely unchanged. This would suggest that ultrasound imaging findings of an intact rotator cuff do not add substantial diagnostic value for ruling-in or ruling-out a PAR compared with clinical examination findings alone.

Limitations of the Study

There were some inherent limitations in this study limiting its widespread clinical application in its current form. In designing this study, while the musculoskeletal radiologists and sonographers reached general agreement upon broad diagnostic criteria for ACJ pathology, imaging of the ACJ on ultrasound is a recent development in musculoskeletal sonography, and there was sufficient scope for individual interpretation to make standardisation of definitions difficult. Hence consensus on strictly standardised ACJ diagnostic criteria was not reached prior to commencing the study, and some retrospective collection of clinical data regarding ACJ pathology was used for this analysis. The validity of ultrasound for detecting ACJ pathology also requires further investigation by comparing findings with MRI or surgical observations. These results need to be verified in a larger sample of participants using more standardised definitions of ACJ pathology. The false-positive rate of ACJ diagnostic blocks has not been reported and results should be interpreted in this context.

Conclusion

In conclusion, these findings represent preliminary results that ACJ capsular hypertrophy identified on diagnostic ultrasound may be a useful indicator of symptomatic ACJ pathology defined by an 80% PAR. When added to individual and clinical test combinations, this imaging finding considerably improved the ability to predict a PAR, however clinical examination findings alone were sufficient to rule-out a PAR. Pending further validation of reported ACJ pathology on diagnostic ultrasound, these results may guide the differential diagnosis of painful ACJ conditions from other sources of shoulder pain, and provide rationale for the use of diagnostic ultrasound

imaging for this purpose. These combinations of clinical and imaging features may also provide an indication of those likely to respond to targeted ACJ pain relief interventions, or who may be appropriate for more expensive or invasive investigation or treatment procedures.

CHAPTER SEVEN

CLINICAL DIAGNOSIS OF ROTATOR CUFF TEARS

Preface

This chapter relates to Specific Aim 5 of the thesis:

To estimate the diagnostic accuracy of clinical examination findings for identifying large rotator cuff tears.

Large rotator cuff tears are of diagnostic and prognostic significance, given their potential to alter management that frequently involves surgical intervention. This final chapter of the thesis presents diagnostic accuracy results for the ability of clinical examination findings to detect the presence of a large or multi-tendon rotator cuff tear identified by diagnostic ultrasound scan.

These results have been presented in the form of a manuscript that will be submitted to a peer-reviewed journal upon completion of this thesis. To avoid repetition of methodology that has already been presented, the clinical examination and diagnostic ultrasound procedures that have previously been described are not repeated in detail, however descriptions of these procedures can be found in Chapter 4.

7.1. DIAGNOSTIC ACCURACY OF THE CLINICAL EXAMINATION FOR IDENTIFYING LARGE ROTATOR CUFF TEARS

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Taylor, S. Diagnostic accuracy of the clinical examination for identifying large rotator cuff tears.

The results presented in the following manuscript will be submitted for journal publication upon completion of the PhD.

Abstract

Background: Significant disruption of the rotator cuff can result in considerable pain and loss of function, with surgical management demonstrating superior outcomes for pain and function compared with conservative management. The aim of this study was to identify clinical predictors of medium, large and multi-tendon (MLM) rotator cuff tears to facilitate early identification and referral for orthopaedic evaluation.

Methods: Consecutive patients with shoulder pain (n=203) underwent a standardised clinical examination (index tests) followed by a diagnostic ultrasound scan (reference standard test). A multivariate prediction model was derived with the strongest predictive ability for a MLM rotator cuff tear. Diagnostic accuracy of the prediction variables was compared with diagnostic accuracy of combinations of clinical variables associated with PAR to SAB diagnostic block.

Results: A MLM rotator cuff tear was identified in 12% of participants. Constant pain (AOR 3.04; 95% CI 1.11, 8.30) and painful arc during abduction (AOR 13.97; 95% CI 1.81, 108.82) were the strongest predictors of a MLM rotator cuff tear. Speed's test was the most sensitive (0.96; -LR 0.21) and external rotation lag sign the most specific (0.97; +LR 4.43) of the individual clinical variables for a MLM rotator cuff tear. Using combinations of clinical examination findings, less than five variables demonstrated 100% sensitivity (-LR 0.00), and eight or more variables demonstrated 91% to 100% specificity (+LR >4.66) for a MLM rotator cuff tear.

Conclusion: Several individual tests and combinations of clinical examination findings may assist the clinician in identifying patients with a MLM rotator cuff tear for whom early referral for further investigation or orthopaedic evaluation may be required.

Introduction

The rotator cuff aids in both movement and stability of the glenohumeral joint, assisting in abduction, external and internal rotation movements of the humerus (McCabe et al., 2005). Coordinated rotator cuff activity centres the humeral head in the glenoid assisting control of multidirectional stability (Boettcher, Ginn, & Cathers, 2008), and reduces subacromial pressure by countering the upwards pull of deltoid during arm elevation (Zingg et al., 2007).

Rotator cuff tears disrupt the normal function of the rotator cuff causing pain and dysfunction. Rotator cuff tears vary in size, symptomatology, and while some tears may be asymptomatic (Milgrom et al., 1995; Tempelhof et al., 1999), large tears are associated with significant weakness and loss of function especially in younger patients (Gerber, Fuchs, & Hodler, 2000; Itoi, Minagawa, Sato, Sato, & Tabata, 1997; McCabe et al., 2005). While several classification systems for rotator cuff tear size have been proposed, most define a 'small tear' as being less than 10mm in size, 'medium' tears 10-30mm in size, and a 'large' or 'massive' tear as being more than 30mm in size, or with involvement of two or more tendons (Bateman, 1963; Gerber et al., 2000; Patte, 1990; S. J. Snyder, 1993).

Rotator cuff tear size and location are of prognostic significance. Small tears with little or no tendon retraction frequently remain small (Yamaguchi et al., 2001), however large tears as well as articular surface partial-thickness rotator cuff tears frequently increase in size (Yamanaka & Matsumoto, 1994) and may result in retraction of tendon ends, progressive fatty infiltration and superior migration of the humerus with narrowing of the acromiohumeral distance. Such changes render large cuff tears irreparable due to the poor tissue quality and altered mechanics which frequently results in osteoarthrosis and poor functional outcomes (Gerber, Wirth, & Farshad, 2011).

Large rotator cuff tears have been identified as one of several prognostic determinants of a poor outcome of conservative management (Bartolozzi, Andreychik, & Ahmad, 1994; Neri, Chan, & Kwon, 2009). Although the optimal timing for surgical intervention is a contentious issue, there is evidence that surgical repair of full thickness tears results in favourable outcomes for pain, strength and function (Burkhart, Barth, Richards, Zlatkin, & Larsen, 2007; Lähteenmäki, Hiltunen, Virolainen, & Nelimarkka, 2007; Levy et al., 1991). Aggressive repair of partial-thickness rotator cuff lesions in elite athletes or high-demand occupations, particularly in the presence of associated labral or other pathology, has also been advocated (Matava, Purcell, & Rudzki, 2005).

The key parameters involved in surgical decisions are the patient's symptoms, reparability of the lesion, and short- and longer-term functional demands (Gerber et al., 2011; Matava et al., 2005). It is therefore important to identify medium to large size, or multiple reparable tendon tears early, prior to loss of tissue viability and narrowing of the acromiohumeral distance to optimise the outcome of any surgical interventions, and to maximise patient functional outcomes.

Clinical examination tests have demonstrated variable diagnostic accuracy for identification of rotator cuff tears (Hegedus et al., 2008). Studies agree that clinical tests lack accuracy for differentiation between early to advanced stages of rotator cuff tears (Park et al., 2005). However "lag" signs in which the inability to maintain the test position due to loss of muscle integrity have shown consistently high levels of specificity (88% to 98%) for large rotator cuff tears (Barth et al., 2006; Castoldi et al., 2009; Gerber et al., 1996; Hertel et al., 1996; Miller et al., 2008b; Murrell & Walton, 2001; Park et al., 2005; Scheibel et al., 2005; G Walch, Boulahia, Calderone, & Robinson, 1998). Fewer studies have investigated other aspects of the clinical examination including history variables (e.g. night pain) and resisted tests as potential clinical predictors of medium, large or multi-tendon tears (Ebell, 2005; Litaker et al., 2000; Murrell & Walton, 2001; Park et al., 2005). All these studies were conducted in orthopaedic settings. The prevalence of medium, large and multi-tendon rotator cuff tears on orthopaedic waiting lists or surgical schedules is likely to be higher than encountered in primary care populations where the majority of these lesions are first seen. Prevalence is known to affect the generalisation of diagnostic accuracy, particularly predictive values, to other settings in which the prevalence differs (Leeflang, Bossuyt, & Irwig, 2009). The prevalence of such tears has not been previously reported in a primary care cohort, and the diagnostic accuracy of clinical tests for identifying a medium, large or multi-tendon tear in this population remains unknown.

The aim of this study was to report the prevalence of medium, large or multi-tendon tears in a primary care population and to identify individual and combinations of clinical examination tests with the highest level of diagnostic accuracy for these lesions.

Methods

Study Design and Setting

Sampling, recruitment, inclusion and exclusion criteria are the same as the previous two manuscripts, and have been previously described within this thesis (Chapter 4, p103).

Procedures

Clinical examination.

All included participants completed self-report questionnaires consisting of SF-8™ health survey (Ware et al., 2001), Shoulder Pain and Disability Index (SPADI) (Roach et al., 1991) and Fear Avoidance Beliefs Questionnaire (FABQ) (Gordon Waddell et al., 1993). This was followed by a standardised clinical examination (history and physical examination) conducted by an experienced clinician (AC). A full list of clinical examination variables and response criteria are presented in Appendix 7 (p273). Indeterminate results of clinical examination tests were recorded and coded as missing data.

Diagnostic ultrasound scan.

Participants underwent a standardised series of shoulder radiographs (x-ray) consisting of anterior-posterior (AP) views in neutral, external and internal rotation, axial view and outlet view (Anderson et al., 1998), followed by a diagnostic ultrasound scan performed by trained and experienced musculoskeletal sonographers and reported by fellowship trained musculoskeletal radiologists. The diagnostic ultrasound procedure is described in Chapter 4 (p105).

Sonographers and radiologists recorded diagnostic information on a standardised worksheet that included rotator cuff tear classification according to location (intrasubstance, articular or bursal surface), size classification (high-grade (more than 50% of tendon thickness); low grade (less than 50% of tendon thickness); or full thickness tear including retraction) and dimensions (width and height in mm). A "medium, large or multi-tendon" (MLM) rotator cuff tear was defined by a 'high grade' partial thickness tear, full thickness tear, any tear exceeding 10mm, or a tear affecting two or more tendons.

Statistical Methods

The Fisher exact test (dichotomous variables) was used to assess the association between individual demographic, self-report questionnaires and clinical examination variables with a MLM rotator cuff tear using Statistical Package for the Social Sciences (SPSS) version 17.0 (IBM® Corporation 2010). Variables demonstrating univariate association with a rotator cuff tear at the $p \leq .200$ level were included in multiple logistic regression analyses and stepwise backward variable elimination was performed using Akaike's Information Criterion (AIC) (Akaike, 1974) to derive the strongest predictors of a MLM rotator cuff tear. Multiple regression analysis was carried out using "R" statistical software (R Development Core Team, 2010). The goodness of fit for the model was assessed using the Hosmer-Lemeshow test (Hosmer & Lemeshow, 2000).

Diagnostic accuracy statistics including sensitivity, specificity, predictive values, positive likelihood ratios (+LR) and negative likelihood ratios (-LR) and 95% confidence intervals (CI) were calculated to assess the discriminatory ability of individual variables associated with a MLM rotator cuff tear, combinations of clinical variables according to the number of variables present, as well as the diagnostic accuracy of the strongest predictor variables derived from the regression analysis. Area under the receiver operator curve (ROC) was assessed to find the optimal number of clinical tests for identifying a MLM rotator cuff tear. Confidence Interval Analysis software (Bryant, 2000) was used for calculation of diagnostic accuracy statistics.

Results

Results of recruitment (Figure 5.1, p133) and participant demographics (Table 5.1, p132) are the same as presented in Chapter 5 with the exception of one participant whose data did not alter the descriptive characteristics.

A total of 203 participants completed the clinical examination and diagnostic ultrasound scan (Figure 5.1, p133). A medium, large or multi-tendon (MLM) rotator cuff tear was identified in 24 participants (11.8%), with four cases (1.9%) identified in which both a large and MLM tear were reported. Frequency distributions of rotator cuff tear size are presented in Table 7.1, and a description of other pathology identified on ultrasound is presented in Table 7.2.

Table 7.1. *Description of Tears in Individual Rotator Cuff Components*

Tear description	Rotator cuff component	n	%
No tear		151	74.3
Small tear (<10mm) (n=29)	Supraspinatus	24	11.8
	Infraspinatus	1	0.5
	Subscapularis	4	1.9
Medium-large tear (≥10mm) (n=24)	Supraspinatus	19	9.4
	Infraspinatus	2	1.0
	Subscapularis	3	1.5
Multiple tendon tears (n=4)	All three tendons	1	0.5
	Supraspinatus and infraspinatus	1	0.5
	Supraspinatus and subscapularis	2	1.0

Note: Four cases were identified in which both a large tear and a multi-tendon tear were present.

Table 7.2. *Distribution of Other Pathology in Groups With and Without Medium-Large or Multi-Tendon Tears*

Pathology identified on ultrasound	MLM rotator cuff tear group (n=24) % with pathology	No MLM rotator cuff tear group (n=179) % with pathology
SAB pathology	63	27***
Dynamic bursal bunching	74	58
Rotator cuff tendinosis	17	15
supraspinatus	13	14
infraspinatus	0	1
subscapularis	4	2
Calcific tendinopathy	21	25
supraspinatus	4	18
infraspinatus	4	5
subscapularis	21	8
LHB pathology	17	1**
Biceps tendon sheath effusion	33	10**
GHJ effusion	17	2
ACJ pathology	26	26

Abbreviations. MLM, medium, large or multi-tendon tear; SAB, subacromial bursa; LHB, long head of biceps; GHJ, glenohumeral joint; ACJ, acromioclavicular joint.
*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$.

Age and the SPADI pain subscale score, mechanism of injury (trauma) and night pain as well as physical examination findings of pain with resisted tests, symptom provocation during passive ROM external rotation (90° abduction), positive external rotation lag sign and Speed's test were associated with the presence of a MLM rotator cuff tear ($p \leq 0.20$) (Table 7.3). For individual variables sensitivity ranged from 0.13 (external rotation lag sign) to 0.96 (Speed's test) and specificity ranged from 0.22 (pain with resisted abduction or external rotation) to 0.97 (external rotation lag sign) (Table 7.3). The highest +LR (4.4) was observed for external rotation lag sign, and the lowest -LR (0.12) was observed for painful arc during abduction.

Table 7.3. *Diagnostic Accuracy of Individual Clinical Examination Variables for a Medium-Large or Multi-Tendon Rotator Cuff Tear*

Clinical variables	Cell counts				Diagnostic accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)
Age >50 years	12	12	50	129	0.50 (0.31, 0.69)	0.72 (0.65, 0.78)	0.19 (0.11, 0.31)	0.92 (0.86, 0.95)	1.79 (1.07, 2.70)	0.69 (0.43, 0.97)	2.58* (1.09, 6.12)
SPADI (pain >48%)	17	7	97	81	0.71 (0.51, 0.85)	0.46 (0.38, 0.53)	0.15 (0.10, 0.23)	0.92 (0.85, 0.96)	1.30 (0.91, 1.66)	0.64 (0.32, 1.12)	2.03 (0.80, 5.13)
Traumatic onset	16	8	61	118	0.67 (0.47, 0.82)	0.66 (0.59, 0.73)	0.21 (0.13, 0.31)	0.94 (0.88, 0.97)	1.96 (1.31, 2.67)	0.51 (0.27, 0.82)	3.87* (1.57, 9.55)
Night pain	18	5	87	89	0.78 (0.58, 0.90)	0.51 (0.44, 0.58)	0.17 (0.11, 0.26)	0.95 (0.88, 0.98)	1.59 (1.15, 2.00)	0.43 (0.19, 0.84)	3.68* (1.31, 10.36)
Resisted tests – pain (abd or ER)	21	2	138	38	0.91 (0.73, 0.98)	0.22 (0.17, 0.29)	0.13 (0.09, 0.19)	0.95 (0.84, 0.99)	1.17 (0.93, 1.32)	0.40 (0.11, 1.27)	2.89 (0.65, 12.88)
PROM ER(90° abd) symptoms provoked	22	2	130	46	0.92 (0.74, 0.98)	0.27 (0.21, 0.34)	0.14 (0.10, 0.21)	0.96 (0.86, 0.99)	1.25 (1.00, 1.42)	0.31 (0.09, 1.00)	3.89 (0.88, 17.20)
ERLS (positive)	3	21	5	171	0.13 (0.04, 0.31)	0.97 (0.94, 0.99)	0.38 (0.14, 0.69)	0.89 (0.84, 0.93)	4.43 (1.20, 15.40)	0.90 (0.71, 0.99)	4.89 (1.09, 21.93)
Speed's test (positive)	22	1	109	64	0.96 (0.79, 0.99)	0.37 (0.30, 0.44)	0.17 (0.11, 0.24)	0.99 (0.92, 1.00)	1.51 (1.23, 1.74)	0.12 (0.02, 0.58)	12.92*** (1.70, 98.12)

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; SPADI, Shoulder Pain and Disability Index; PROM, passive range of motion; abd, abduction; ER, external rotation; ERLS, external rotation lag sign

[a]any one of three resisted tests (abduction, external rotation or internal rotation) demonstrating ≥40% strength deficit compared with unaffected side.

*p≤0.05; **p≤0.01; ***p≤0.001.

The constant nature of pain and a painful arc during abduction were the strongest predictors of a MLM rotator cuff tear (AOR 3.04; 95% CI 1.11, 8.30 and AOR 13.97; 95% CI 1.81, 108.82 respectively) (Table 7.4). Of the two variables, painful arc during abduction demonstrated the highest sensitivity (0.95) and constant pain demonstrated highest specificity (0.72), with highest +LR observed when both were positively identified (3.10) (Table 7.4).

Diagnostic accuracy results for combinations of all clinical examination variables are presented in Table 7.5. Highest sensitivity (1.00) was observed for up to five positive clinical examination findings and highest specificity (1.00) was observed when all 10 clinical examination findings were present. Highest +LR (infinity) was also observed for 10 clinical examination findings, and lowest -LR was observed when less than five clinical examination findings were present. Area under the ROC curve was 0.838 (0.772, 0.905; p=0.000) and five positive clinical examination findings were identified as the optimal diagnostic point with sensitivity and specificity 0.88 and 0.66 respectively.

Discussion

This study provides the first known report of the prevalence of medium, large or multi-tendon (MLM) rotator cuff tears in a primary care cohort. The diagnostic accuracy of clinical examination variables in identifying symptomatic MLM rotator cuff tears in this population was estimated. Although the prevalence of these tears was low (11.8%), they are of diagnostic and prognostic significance. A positive external rotation lag sign, reports of constant pain combined with a painful arc during abduction, as well as combinations of at least eight other history and physical examination variables were highly specific for identification of a MLM rotator cuff tear. These results may be generalised to the majority of primary contact physiotherapy practices in New Zealand, and potentially to other countries with similar physiotherapy training and service delivery strategies in the primary care setting.

Table 7.4. *Diagnostic Accuracy of Clinical Examination Prediction Model Variables for a Medium-Large or Multi-Tendon Rotator Cuff Tear*

Clinical variables	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)	AOR (95% CI)
		Cell counts						Diagnostic accuracy				
Constant pain	13	11	50	129	0.54 (0.35, 0.72)	0.72 (0.65, 0.78)	0.21 (0.13, 0.32)	0.92 (0.87, 0.96)	1.95 (1.20, 2.88)	0.64 (0.38, 0.91)	3.05* (1.28, 7.26)	3.04* (1.11, 8.30)
Painful arc abduction	18	1	86	71	0.95 (0.75, 0.99)	0.45 (0.37, 0.53)	0.17 (0.11, 0.26)	0.99 (0.93, 1.00)	1.72 (1.34, 2.04)	0.12 (0.02, 0.56)	14.86*** (1.94, 114.06)	13.97* (1.81, 108.82)
Both	9	10	24	133	0.47 (0.27, 0.68)	0.85 (0.78, 0.90)	0.27 (0.15, 0.44)	0.93 (0.88, 0.96)	3.10 (1.62, 5.36)	0.62 (0.37, 0.87)	4.99** (1.84, 13.56)	

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; AOR, adjusted odds ratio

*p≤0.05; **p≤0.01; ***p≤0.001.

Table 7.5. *Diagnostic Accuracy for Combinations of Clinical Variables for a Medium-Large or Multi-Tendon Rotator Cuff Tear*

Clinical variable combinations [a]Number of positive clinical tests	Cell counts				Diagnostic accuracy						
	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	+LR (95% CI)	-LR (95% CI)	OR (95% CI)
1 or more	24	0	178	1	1.00 (0.86, 1.00)	0.01 (0.00, 0.03)	0.12 (0.08, 0.17)	1.00 (0.21, 1.00)	1.01 (1.00, 1.03)	0.00 (0.00, 27.73)	0.88 (0.84, 0.93)
2 or more	24	0	166	13	1.00 (0.86, 1.00)	0.07 (0.04, 0.12)	0.13 (0.09, 0.18)	1.00 (0.77, 1.00)	1.08 (1.07, 1.14)	0.00 (0.00, 1.96)	0.87 (0.83, 0.92)
3 or more	24	0	148	31	1.00 (0.86, 1.00)	0.18 (0.13, 0.24)	0.14 (0.10, 0.20)	1.00 (0.89, 1.00)	1.21 (1.18, 1.31)	0.00 (0.00, 0.81)	0.86* (0.81, 0.91)
4 or more	24	0	125	54	1.00 (0.86, 1.00)	0.31 (0.24, 0.38)	0.16 (0.11, 0.23)	1.00 (0.94, 1.00)	1.43 (1.36, 1.59)	0.00 (0.00, 0.46)	0.84*** (0.78, 0.90)
5 or more	24	0	90	89	1.00 (0.86, 1.00)	0.50 (0.43, 0.57)	0.21 (0.15, 0.29)	1.00 (0.96, 1.00)	1.99 (1.78, 2.32)	0.00 (0.00, 0.28)	0.79*** (0.72, 0.87)
6 or more	21	3	61	118	0.88 (0.69, 0.96)	0.66 (0.59, 0.73)	0.26 (0.17, 0.36)	0.98 (0.93, 0.99)	2.57 (1.91, 3.27)	0.19 (0.07, 0.47)	13.54*** (3.89, 47.20)
7 or more	15	9	35	144	0.63 (0.43, 0.79)	0.81 (0.74 0.86)	0.30 (0.19, 0.44)	0.94 (0.89, 0.97)	3.20 (2.08, 4.91)	0.47 (0.28, 0.79)	6.86*** (2.77, 16.95)
8 or more	10	14	16	163	0.42 (0.25, 0.61)	0.91 (0.86, 0.95)	0.39 (0.22, 0.58)	0.92 (0.87, 0.95)	4.66 (2.34, 8.74)	0.64 (0.43, 0.83)	7.28*** (2.79, 19.01)
9 or more	5	19	3	176	0.21 (0.09, 0.41)	0.98 (0.95, 0.99)	0.63 (0.31, 0.86)	0.91 (0.85, 0.94)	12.43 (3.40, 44.18)	0.81 (0.61, 0.93)	15.44*** (3.42, 69.72)
10	2	22	0	179	0.08 (0.02, 0.26)	1.00 (0.98, 1.00)	1.00 (0.34, 1.00)	0.89 (0.84, 0.93)	~ (1.78, 728.00)[†]	0.92 (0.79, 1.03)[†]	9.14* (6.16, 13.55)

Abbreviations. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; OR, odds ratio; ~, infinity
[a]refers to combinations of clinical tests including any of: age >50 years, SPADI (pain) score >48%, traumatic onset, constant pain, night pain, painful arc abduction, pain with resisted abduction or external rotation, symptoms reproduced with passive external rotation (at 90° abduction), positive external rotation lag sign, positive Speed's test
*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$.
[†]0.5 added to cells to estimate 95% confidence intervals.

Other studies report increases in prevalence of asymptomatic rotator cuff tears with advancing age, becoming particularly prevalent in those over 50 to 60 years of age (Milgrom et al., 1995; Sher et al., 1995; Tempelhof et al., 1999). Results from the current study involving symptomatic participants also identified a relationship between increasing age and medium-large or multi-tendon cuff tears, with those over the age of 50 years being more than twice as likely to be diagnosed with a significant rotator cuff tear, however age alone was not sufficient to discriminate between those with and without a significant rotator cuff tear.

Individual Clinical Examination Tests

Our results support previous findings in which a positive external rotation lag sign (ERLS) was reported to be highly specific for a rotator cuff tear (94% to 100%) (Castoldi et al., 2009; Hertel et al., 1996; Miller et al., 2008a). However in contrast to previous studies in which sensitivity was also reported to be high (0.94), sensitivity of this test in the current study was low (0.13). This is likely to be due to the lower prevalence of significant rotator cuff tears in the primary care population compared to orthopaedic settings in which the previous studies were conducted. Overall, specificity values were variable (0.22 to 0.97), suggesting a number of other structures other than the rotator cuff may also present with similar signs and symptoms, and indeed a number of other pathologies were identified on ultrasound in the group with a significant rotator cuff tear, particularly subacromial bursal pathology and bunching.

The strongest predictors of a medium-large or multi-tendon cuff tear were constant pain and a painful arc during abduction. A painful arc of abduction has long been used in the diagnosis of 'impingement' which includes rotator cuff tears (Cyriax, 1978) and has also previously been identified as a strong predictor of supraspinatus pathology on ultrasound (Chew et al., 2004; Moosikasuwan, Miller, & Burke, 2005). The combination of constant pain and painful arc demonstrated 85% specificity (lower 95% confidence limit of 0.78), and a +LR of 3.10 representing moderate to high levels of confidence for ruling-in the diagnosis of a significant rotator cuff tear. Although the painful arc abduction also demonstrated 95% sensitivity for a large rotator cuff tear, there were a number of indeterminate results for this test. Most indeterminate results were related to insufficient ROM of abduction in participants with shoulder stiffness or high levels of pain severity affecting the ability to raise the arm to the required level of abduction to complete the test. This may have introduced bias into sensitivity estimates and this may limit the clinical application of this test when used in isolation.

Pain provocation during Speed's test (resisted straight-arm raise), resisted abduction or external rotation, passive external rotation at 90° abduction and a painful arc during abduction all demonstrated in excess of 90% sensitivity, and apparent ability to rule-out a significant rotator cuff tear. Speed's test demonstrated the highest sensitivity (0.96) and low -LR (0.00) enabling a significant cuff tear to be ruled-out when this test is negative in this sample. Speed's test is reported to predominantly stress the long head of biceps tendon (Holtby & Razmjou, 2004), however the intimate anatomic connections between the biceps tendon, rotator interval, subscapularis and supraspinatus tendons and association between anterior-superior rotator cuff tears and rotator interval injury (Gaskill, Braun, & Millett) mean injury to any of these structures may result in pain provocation during this test. This may explain the high sensitivity but low specificity of Speed's test for a MLM rotator cuff tear. Although the sensitivity of resisted abduction or external rotation, and passive external rotation was high and the -LRs were small (0.40 and 0.31 respectively), the upper 95% confidence limit reached 1.00, reducing confidence in the likelihood of no significant tear not being present when these tests are negative.

Combinations of Clinical Examination Features

Combinations of ten history and physical examination variables demonstrated 100% sensitivity for a medium-large or multi-tendon rotator cuff tear with almost zero-odds of a significant rotator cuff tear when fewer than five of the tests were positive. The same combination of ten tests ruled-in a significant rotator cuff tear with more than 80% specificity when at least seven tests were positive, with 25% of participants fitting this criterion. Those with at least eight positive tests were almost five times more likely to have a significant rotator cuff tear (specificity 0.91), increasing to 12 times more likely when nine tests were positive (specificity 0.98). Ten positive tests resulted in 100% specificity and a +LR of infinity, however only two participants fitted this criterion.

Limitations of the Study

Limitations included the potential for diagnostic ultrasound to miss subtle partial-thickness articular surface rotator cuff tears, the aim was to identify only medium-large rotator cuff tears, for which ultrasound sensitivity is known to be higher than for small tears (Fotiadou et al., 2008; Iannotti et al., 2005). In addition, high-frequency

transducers and high resolution monitors were used, and all staff were trained and experienced musculoskeletal sonographers and radiologists.

Conclusion

Although the prevalence of significant rotator cuff tears is low in this primary care cohort, these medium, large and multi-tendon rotator cuff tears have significant implications for referral and surgical management for improvements in patient outcome. Such lesions can be identified using individual and combinations of clinical tests. When these tests are positive, early referral for additional imaging to determine the extent or associated pathology, or referral for orthopaedic consultation may be warranted.

CHAPTER EIGHT

SUMMARY, DISCUSSION AND CONCLUSIONS

Preface

The final chapter of this thesis begins with a summary of the main findings from individual thesis chapters. This is followed by a general discussion of the main findings, including study limitations. The clinical applications of the results are discussed and areas of ongoing analysis and future research recommendations are outlined. Finally, the thesis conclusions are presented.

8.1. SUMMARY OF DIAGNOSTIC ACCURACY STUDY RESULTS

The diagnostic accuracy study involved a clinical examination followed by a standardised x-ray series, diagnostic ultrasound scan, and diagnostic blocks into the SAB and ACJ. Those not reporting at least 80% pain relief following the SAB or ACJ diagnostic block also received a diagnostic block into the GHJ performed as part of an MR arthrogram procedure. In Chapter 4, the prevalence of imaged pathology was reported, and the relationship between imaging findings and response to diagnostic blocks was evaluated. In Chapters 5 and 6, the diagnostic accuracy results for the ability of the clinical examination to predict a PAR following SAB and ACJ diagnostic block were presented, and the added value of imaging findings for predicting a PAR was reported. A stated aim of the thesis was to also assess the diagnostic accuracy of clinical examination and imaging findings for predicting a PAR following a GHJ diagnostic block. However this analysis was beyond the scope of this thesis, and analysis of this data will be undertaken following thesis submission with results to be reported in due course. In Chapter 7 the clinical examination predictors of a large or multi-tendon rotator cuff tear were identified and diagnostic accuracy of these variables was estimated.

Prevalence of Imaged Pathology and Response to Diagnostic Blocks

Chapter 4 presented the prevalence of imaged pathology on x-ray, ultrasound and MRA in a primary care cohort, as well as reporting the relationship between imaged pathology and response to diagnostic blocks into the SAB, ACJ and GHJ. The most frequent pathologies identified on x-ray were ACJ and GHJ pathology reported in 34 of 203 participants (17%). Rotator cuff pathology was the most common ultrasound finding reported in 102 of 203 participants (50%) with rotator cuff tears the most common rotator cuff pathology, reported in 53 of 203 cases (26%) (Figure 4.5, p114). Multiple pathology was common on MRI, with SAB pathology reported in 71 of the 93 participants (76%), and rotator cuff pathology, GHJ pathology or ACJ pathology reported in 60% or more of the 93 participants who received this investigation.

A small number of x-ray and ultrasound imaging findings were strongly associated with PAR following SAB diagnostic block (x-ray and ultrasound evidence of supraspinatus calcific tendinopathy) and GHJ diagnostic block (absence of rotator cuff tear and biceps tendon sheath effusion on ultrasound). Several ultrasound imaging variables were also associated with a PAR following ACJ diagnostic block (ACJ pathology and absence of a supraspinatus articular surface tear). No MRI variables were strongly associated with anaesthetic responses. The value of these imaging findings for aiding identification of a predominant SAB or ACJ pain source when combined with clinical examination findings was reported in Chapters 5 and 6.

Predictors of Response to Diagnostic Blocks

Subacromial Bursa Diagnostic Block

In Chapter 5, two manuscripts were presented. One manuscript reported diagnostic accuracy results of clinical examination findings for predicting a PAR following SAB diagnostic block, and the other reported the added value of imaging findings for predicting a PAR. A summary of how both sets of results may be combined into a diagnostic reasoning pathway is presented in Figure 8.1.

This pathway includes the two sets of clinical examination combinations. One is a combination of the three strongest predictors of a PAR (Table 5.4, p136), however this combination was unable to rule-out a PAR with any degree of certainty. The second combination of nine clinical examination variables is also included that demonstrated an ability to both rule-in and rule-out a PAR (Table 5.5, p138). When only a small number of these clinical tests were present, identification of supraspinatus calcification or full

thickness tear improved the ability to rule-in an 80% PAR to SAB diagnostic block for a small number of participants. For those who did not satisfy clinical examination or imaging criteria with high levels of sensitivity or specificity for a PAR, a diagnostic injection of local anaesthetic may be the most efficient method of identifying a predominant subacromial source of pain.
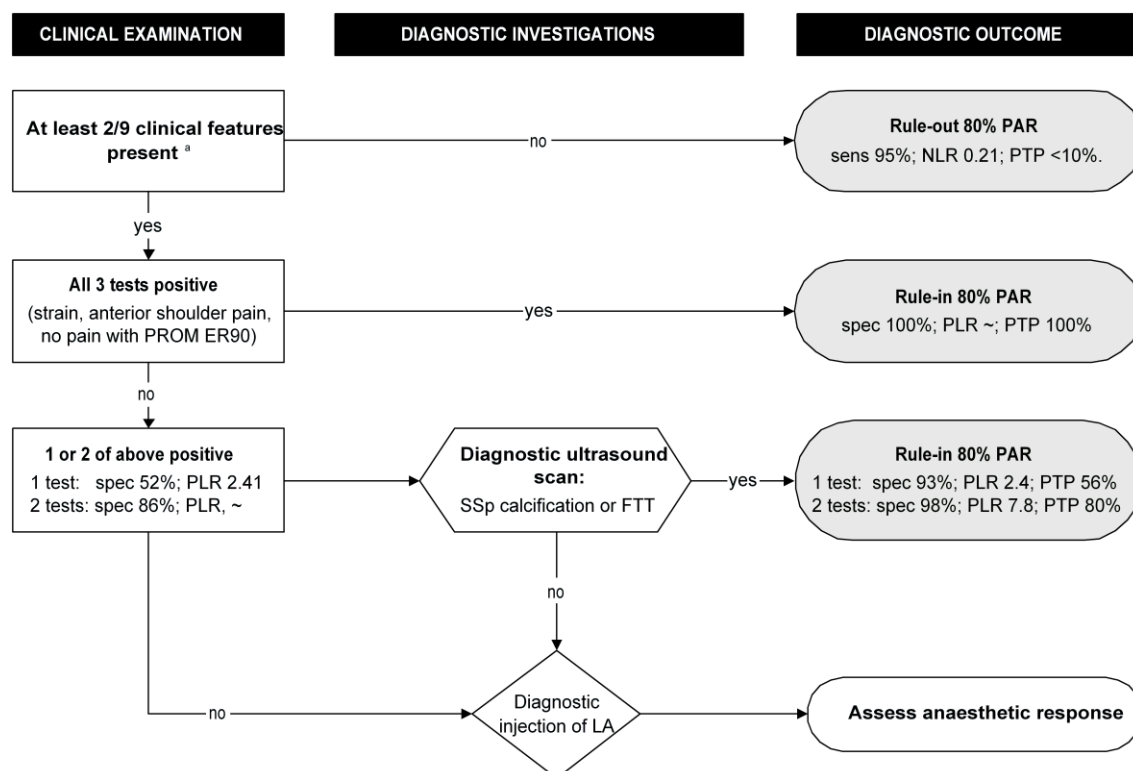

Figure 8.1. Clinical examination and diagnostic imaging pathway for positive response to SAB diagnostic block. [a]Any test combination including strain injury, anterior shoulder pain, inability to sleep on affected side, hand-behind back asymptomatic, passive GHJ abduction, external rotation (90° abduction), internal rotation (90° abduction), cross body adduction asymptomatic, negative Hawkins-Kennedy test. *Abbreviations.* PAR, positive anaesthetic response; sens, sensitivity; -LR, negative likelihood ratios; PTP, post-test probability for a PAR (based on a pre-test probability of 34%); PROM ER90, passive range of motion external rotation (at 90° abduction); spec, specificity; +LR, positive likelihood ratio; ~, infinity; SSp, supraspinatus; FTT, full-thickness tear; LA, local anaesthetic.

Acromioclavicular Joint Diagnostic Block

In Chapter 6.1, two individual clinical examination features (Table 6.3, p167), and combinations of nine variables (Table 6.4, p168) were identified with the ability to rule-in and rule-out a PAR. In Chapter 6.2, five of the original nine clinical variables were also identified that improved the post-test probability for a PAR compared with the combinations of the nine clinical features (Table 6.7, p178). Pending further research, preliminary results indicate that the identification of ACJ pathology on ultrasound, and particularly the presence of capsular hypertrophy may considerably improve the

probability of an ACJ PAR. When the source of pain remains unclear, additional diagnostic tests may be required such as a diagnostic injection or MRI scan. These results are summarised into a diagnostic reasoning pathway presented in Figure 8.2.
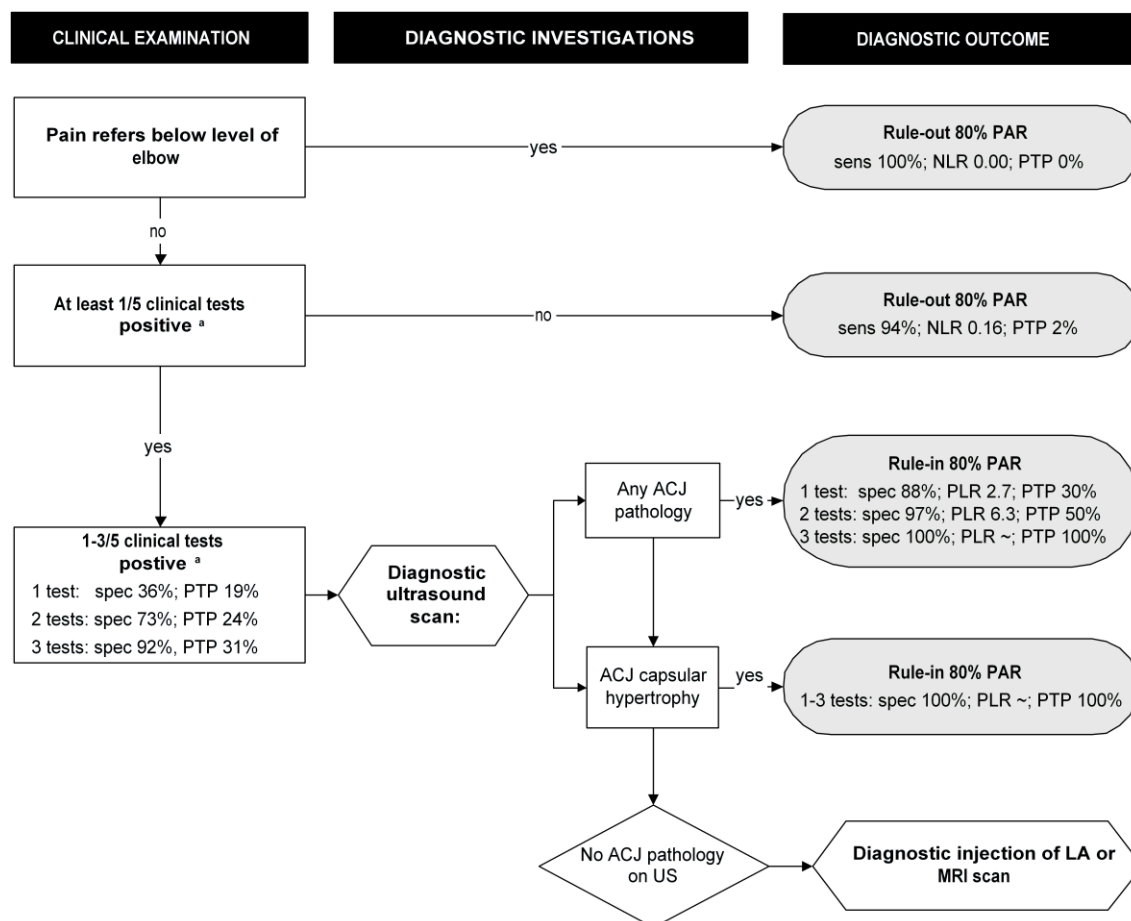

Figure 8.2. Clinical examination and diagnostic imaging pathway for positive response to ACJ diagnostic block. [a]Any test combination including repetitive onset of pain, absence of referred pain below the level of the elbow, thickened or swollen ACJ, no symptom provocation during passive GHJ abduction or external rotation (at 90° abduction). *Abbreviations.* PAR, positive anaesthetic response; sens, sensitivity; -LR, negative likelihood ratios; PTP, post-test probability for a PAR (based on a pre-test probability of 14%); spec, specificity; +LR, positive likelihood ratio; ~, infinity; ACJ, acromioclavicular joint; US, diagnostic ultrasound; LA, local anaesthetic; MRI, magnetic resonance imaging.

Clinical Predictors of Large Rotator Cuff Tears

The prevalence of medium, large or multi-tendon rotator cuff tears was low in this primary care cohort (11.8%) however, these lesions were accurately identified using individual, and combinations of clinical tests (Tables 7.3 and 7.4, p194-196). When clinical findings indicate the presence of these lesions, early referral for additional imaging may be appropriate to determine the extent of the lesion or the presence of

associated pathology, or referral for orthopaedic consultation may be required. Additional diagnostic investigations (diagnostic ultrasound) scan may be justified when results of clinical tests are unclear. The diagnostic reasoning pathway and associated diagnostic estimates for each step in this process are presented in Figure 8.3.
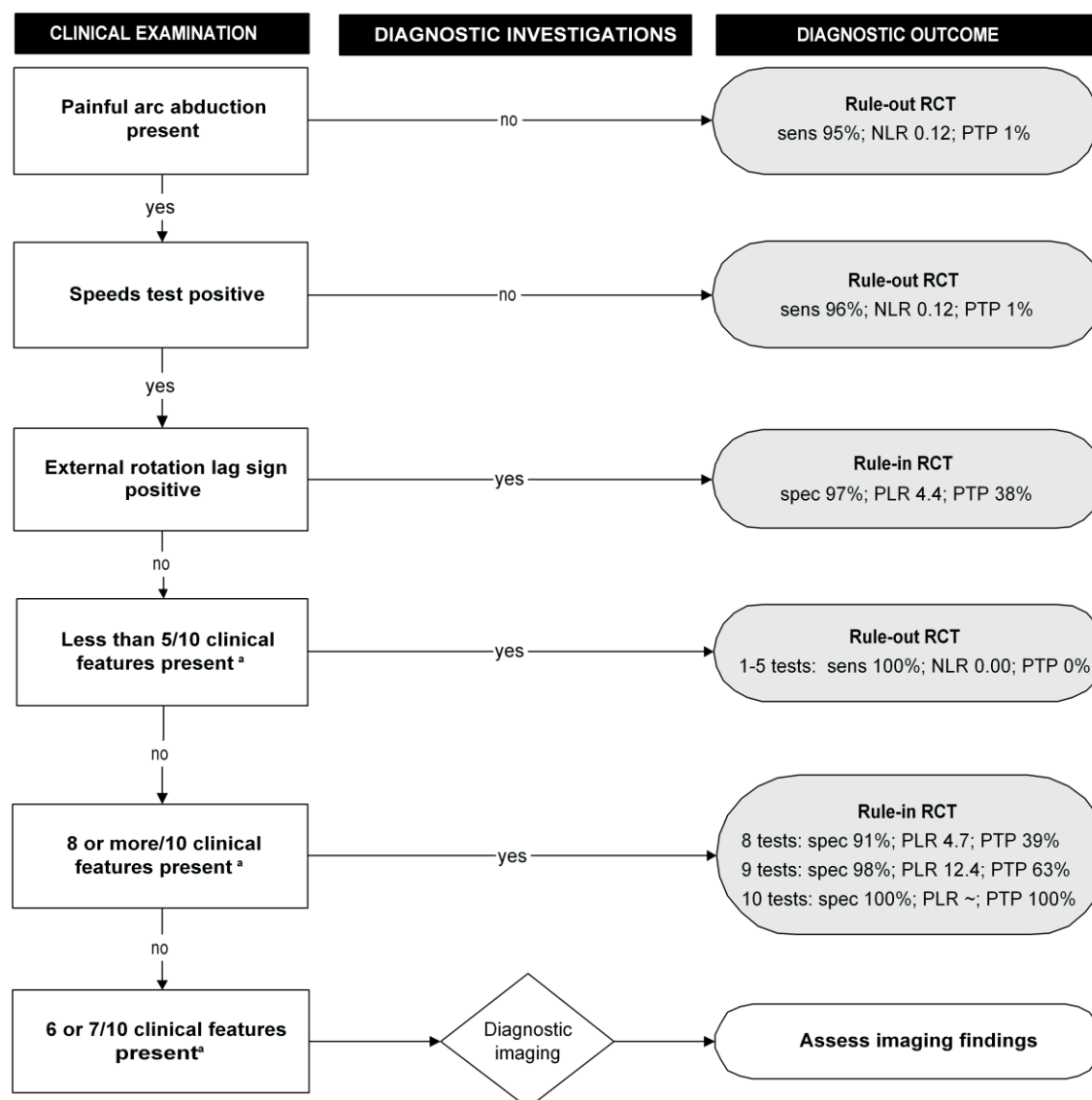

Figure 8.3. Clinical examination pathway for medium-large or multi-tendon rotator cuff tear. [a]Any test combination including age >50 years, SPADI pain score >48%, traumatic onset of pain, constant pain, night pain disturbs sleep, painful arc abduction present, symptoms provoked with resisted abduction or external rotation, symptoms provoked with passive external rotation (at 90° abduction), positive external rotation lag sign, positive Speed's test. *Abbreviations.* RCT, medium-large or multi-tendon rotator cuff tear; sens, sensitivity; -LR, negative likelihood ratios; PTP, post-test probability for a medium, large or multi-tendon rotator cuff tear (based on pre-test probability of 11.8%); spec, specificity; +LR, positive likelihood ratio; ~, infinity.

.

8.2. DISCUSSION

Diagnostic tests should be considered for use if they yield clinically important information where results may potentially alter patient management and there is a reasonable probability that this altered management will lead to an improvement in patient outcome (Sackett, 1992; Thornbury, 1994). The diagnosis of painful shoulder conditions is challenging, and despite the importance of the clinical examination for informing diagnostic reasoning and treatment selection, no evidence was available for the diagnostic accuracy of the clinical examination for identifying common sources of shoulder pain using response to injection of local anaesthetic as the reference standard. Despite frequent use of diagnostic imaging investigations to aid identification of the likely source of symptoms, the relationship between imaging findings and symptoms was unclear, and the impact of imaging findings on the diagnosis of painful shoulder conditions was unknown. Although the diagnostic accuracy of physical examination tests for identifying the presence of a rotator cuff tear had been extensively reported in patients awaiting surgery, the accuracy of these tests had not been reported in a primary care cohort. This thesis aimed to address these issues.

Positive Response to Diagnostic Blocks

Diagnostic blocks are the accepted reference standard for identifying the tissue source of pain (Bogduk, 2004a), and a positive anaesthetic response (PAR) was defined as a reduction in pain intensity of 80% or more, based on pre- and post-injection provocative clinical testing. While this does not definitively exclude the targeted structure as the source of pain when the anaesthetic response is less than 80%, for the purposes of diagnosis, this provides confidence that the targeted structure was the predominant source of pain.

Identification of the predominant source of pain may help to differentiate involvement of different shoulder structures for which specific management pathways exist or additional medical, radiological or surgical investigations may be required. Painful conditions affecting the SAB include primary synovitis (bursitis), calcific lesions, or may be secondary to repeated mechanical 'impingement' against the acromial arch (Farin et al., 1990; Neer, 1983; Salzman et al., 1997). Pain arising from the ACJ may result from instability, degenerative or post-traumatic arthropathy, inflammatory arthropathy, crystal arthropathy and osteolysis (Deitch, 2004; Johansen et

al., 2011; Mazzocca et al., 2007; A. P. Wright et al., 2011). Identification of these structures as the predominant pain source is an important first step, following which additional investigations such as diagnostic imaging or laboratory testing may be required to differentiate specific pathology and identify additional appropriate management interventions.

The ability to accurately identify (rule-in) those who are likely to report a PAR following a diagnostic block may aid differentiation of the targeted structure from other sources of shoulder pain, and may also provide an indication of the likely therapeutic benefit of targeted pain relief interventions such as corticosteroid injections that are commonly used to treat subacromial bursitis and ACJ conditions (Skedros et al., 2007). The ability to identify those patients who are unlikely to report (rule-out) an 80% PAR following diagnostic block may provide information that enables diagnostic consideration of co-existing or alternate sources of shoulder pain, and also may provide an indication of those who are unlikely to gain significant benefit from targeted pain relief interventions.

While the 80% anaesthetic response criterion provides a high level of confidence for diagnostic purposes of ruling-in or ruling-out a predominant subacromial or ACJ pain source, reports of reductions in pain intensity of lower magnitudes such as 60% or 70% may still represent therapeutically beneficial outcomes following pain relief interventions. Ongoing analyses will evaluate clinical examination and imaging predictors across a range of anaesthetic response criteria, and further research may help determine the ability of anaesthetic response criteria such as 60% or 70% to predict the outcome of specific pain-relief interventions such as corticosteroid injections.

Diagnostic Accuracy of the Clinical Examination for a Positive Anaesthetic Response

The main findings from the results presented in Chapters 5 and 6 provide evidence for the importance of the combination of both history and physical examination findings in identifying a predominant subacromial or ACJ pain source. Several aspects of patient history were identified that were independently associated with an 80% PAR following SAB and ACJ diagnostic blocks. Mechanism of pain onset and location of pain were identified as the strongest clinical predictors of a SAB PAR (strain injury and anterior shoulder pain), and were the most sensitive (no referral of pain below the elbow) and specific (repetitive mechanism of injury) clinical variables for an 80% PAR following ACJ diagnostic block. Few studies had previously investigated the diagnostic accuracy

of specific aspects of patient history, and our results highlight the importance of this commonly acquired clinical information in the diagnostic process. When these history variables were combined with other history and physical examination features they contributed to the ability to rule-out a PAR with high levels of precision for both the SAB and ACJ diagnostic block.

From the physical examination, resisted tests, orthopaedic tests, and other aspects of the physical examination that were previously reported to be either sensitive or specific for subacromial or ACJ pain in orthopaedic settings, including the Hawkins-Kennedy test (Calis et al., 2000), palpation of the ACJ (Walton et al., 2004) and active compression test (O'Brien et al., 1998), were of limited diagnostic value for predicting a PAR following SAB or ACJ diagnostic block in this primary care study. These results support previous reports of low specificity of many physical examination tests in which provocation of pain is used as the criterion for a positive result (MacDonald et al., 2000; Park et al., 2005). Previous studies almost exclusively used surgical or imaging reference standards tests in which identified pathology was assumed to be painful and of symptomatic relevance. Although we used a reference standard specifically designed to identify painful tissue, the specificity of individual tests was still shown to be poor. The complex anatomy of the shoulder region is a likely explanation for the low specificity as it is difficult to stress intricately connected anatomical structures in isolation.

The results of this study did however, identify an interesting trend. While pain provocation with physical examination tests did not appear to be specific for a PAR, the absence of symptom provocation associated with three passive ROM tests used in this study consistently demonstrated an association with, and moderate to high levels of specificity (0.62 to 0.86) for PAR following both the SAB and ACJ diagnostic blocks. Those in whom symptoms were not provoked during passive GHJ abduction, passive external rotation (90° abduction) and passive cross-body adduction were more likely to report a PAR to both the SAB and ACJ diagnostic blocks.

All three of these tests, although also used as diagnostic tests for a variety of other shoulder conditions, involve stress on a number of capsuloligamentous and intra-articular GHJ structures (Calis et al., 2000; Cyriax, 1978; Hegedus et al., 2008; F. W. Jobe & Kvitne, 1989; S. Kim et al., 1999; S. H. Kim et al., 2001; Lo, Nonweiler, Woolfrey, Litchfield, & Kirkley, 2004; Mimori et al., 1999; W. E. Palmer & Caslowitz, 1995; Parentis et al., 2006). The association between the absence of pain with these tests and a PAR may suggest the absence of symptomatic GHJ pathology to be a prerequisite

for reporting a PAR following SAB or ACJ diagnostic block. It is possible that the ability to accurately diagnose painful SAB or ACJ disorders may improve when tests that are provocative of symptomatic GHJ pathology such as external rotation, abduction and cross-body adduction are negative. Pending further interexaminer reliability testing, these tests may warrant further investigation as being potentially valuable for differentiating GHJ structures from other sources of shoulder pain.

Diagnostic Accuracy of Individual Tests Versus Combinations of Clinical Examination Features

Combinations of clinical examination findings demonstrated consistently higher levels of precision of estimates (narrow confidence intervals), and resulted in larger changes in post-test probabilities for PAR following diagnostic block compared with individual clinical examination tests alone. This may be related to complexities associated with measuring the outcome of pain, and specifically a change in pain intensity, compared with an observable and measureable outcome that can be consistently visualised on imaging, such as a rotator cuff tear. Pain is a complex physiological phenomenon that cannot be directly measured, nor observed on imaging. Perception of pain is affected by a complex interaction of a number of factors including nociceptive stimulation, attitudes and beliefs, psychological distress, illness behaviour and social factors (G. Waddell, Bircher, Finlayson, & Main, 1984). Such complex interaction of multiple factors may render it difficult to make clinical decisions based upon change in pain intensity following an isolated test procedure.

The criteria that included combinations of clinical tests with high levels of sensitivity and specificity were defined by a minimum number of positive clinical findings, rather than requiring specific combinations of individual test findings to be present. Using combinations of minimum numbers of positive tests (one or more, two or more, etc), a trend of decreasing sensitivity and increasing specificity was observed. This appears to have been primarily due to an increased proportion of false negatives and true negatives as the required number of positive clinical examination features increased (fewer cases fitting higher test criteria) (Sackett, 1983). These results support the findings of other studies in which sensitivity was higher when fewer positive clinical features were required (Calis et al., 2000; Chronopoulos et al., 2004; Michener et al., 2009; Walsworth et al., 2008). Specificity has also been shown to increase with increasing numbers of clinical features for the diagnosis of rotator cuff tears (Murrell & Walton, 2001), response to lumbar provocation discography (Laslett, Aprill, McDonald,

& Oberg, 2006) and a PAR following sacroiliac joint (Laslett et al., 2005) and lumbar zygapophyseal joint diagnostic blocks (Laslett, McDonald, et al., 2006).

Additional Value of Imaging Findings for Predicting a Positive Anaesthetic Response

Few imaging variables were identified that were associated with either a SAB or ACJ PAR in this study. Ultrasound evidence of supraspinatus calcification or a full-thickness tear were associated with a PAR following SAB diagnostic block, and preliminary results also suggest any ACJ pathology, and particularly ACJ capsular hypertrophy are related to a PAR following ACJ diagnostic block. The low number of imaging findings that were associated with the PAR may be explained by the low prevalence of some of the imaging findings in this primary care study. This may have resulted in larger p-values, and wide 95% confidence intervals for odds ratios for imaging variables, and further investigation in larger samples of primary care patients would help to verify these results. It is also possible that the use of different anaesthetic response criteria (60%, 70%, 90% etc.) may yield other imaging predictors of a PAR that may still represent clinically meaningful diagnostic information with respect to symptomatic shoulder pathology.

When these imaging findings were added to clinical examination features found to be associated with a PAR a consistent trend was observed for both the SAB and ACJ diagnostic blocks. In cases where higher numbers of clinical examination features were present, the addition of imaging findings made little difference to the post-test probability for a PAR. However when few clinical features were present the imaging findings resulted in a substantial increase in the post-test probability of a PAR following both the SAB and ACJ diagnostic block.

When few clinical features were present, confirmation of the presence of these pathologies on ultrasound increased the probability of a PAR following both the SAB and ACJ diagnostic block, resulting in increased confidence that the observed lesions were of symptomatic significance. However the addition of imaging findings to the clinical examination findings only resulted in identification of a small additional number of participants who were likely to report a PAR. This is most likely due to the low prevalence of these specific imaging findings. When ultrasound findings of supraspinatus pathology were added to clinical test combinations, only 5% to 10% of participants satisfied the clinical criteria that demonstrated high levels of specificity for a SAB PAR, and when imaged ACJ pathology was added to clinical test combinations,

only 4% to 13% of participants satisfied criteria with high specificity for an ACJ PAR. In clinical practice, the additional value of the imaging findings would therefore need to be considered against the availability and cost of the imaging procedures.

The addition of imaging findings to clinical test combinations did not improve the ability to rule-out a PAR The reduction in the number of participants who fitted the criteria that included imaging findings (true positives) had an adverse effect on sensitivity in many cases. When imaging findings were added to clinical criteria for an 80% PAR, sensitivity reduced, -LRs approached 1.0, and NPV ranged between 68% and 97% suggesting the clinical examination features alone were better able to rule-out a PAR compared with combinations of clinical and imaging information. The high false-negative rate meant that between 67% and 94% of those who later reported a PAR did not fit the combined clinical and imaging criteria. The clinical implication is that while combinations of clinical and imaging information may identify a small proportion of those likely to report a PAR, in the majority of cases a PAR could not be ruled-out using combinations of clinical and imaging information and additional diagnostic investigations may be required to positive identify those with a predominant subacromial or ACJ pain source.

Additional Diagnostic Investigations

For those who do not satisfy the clinical examination and imaging criteria found to be highly specific for a PAR following SAB and ACJ diagnostic block, additional diagnostic procedures may be considered to aid in ruling-in or ruling-out a predominant subacromial or ACJ pain source (80% PAR). The decision to utilise such additional diagnostic procedures is guided by the cost and availability of the investigation (MRI), the availability of appropriately trained and skilled practitioners, as well as the probability that information gained from these investigations will inform management decisions and improve patient outcomes.

Diagnostic injections provide a method of assessing the relative contribution of the injected structure to symptoms. Diagnostic injections are minimally invasive procedures and are considered to be an important tool to help clarify challenging situations in which the source of pain is unclear (Bogduk, 2004b). Injection accuracy for SAB injections has been shown to approach that of imaging-guided procedures when performed by experienced practitioners (Rutten et al., 2007). Injections into the ACJ are reported to be more technically challenging due to its small size and the frequent presence of osteophytes and anatomic variations, with lower accuracy rates (40% to

67%) of blind injections reported (Bisbinas et al., 2006; Partington & Broome, 1998). Although they have been advocated for many years to assist in the diagnosis of shoulder pain (Cyriax, 1978; Neer, 1972), these procedures are rarely performed in the clinical setting. Given the low prevalence of imaging findings that were associated with an 80% PAR in this study, when factors such as cost and availability of diagnostic imaging are considered, diagnostic injections may represent a cost-effective alternative to diagnostic imaging for identification of a predominant subacromial or ACJ pain source.

Magnetic resonance imaging is another diagnostic tool that provides improved visualisation of pathologies such as glenoid labral lesions, tendon pathology and bony oedema (Shahabpour et al., 2008). Magnetic resonance features, including capsular hypertrophy, joint effusion, advanced degenerative changes and reactive bone oedema in the lateral clavicle or acromion, have been reported to be associated with symptomatic conditions affecting the ACJ identified by response to injection of local anaesthetic (Shubin Stein et al., 2006; Strobel et al., 2003). Caudal osteophytes were reported to be the most sensitive for a PAR (82%) and subchondral bone cysts the most specific (97%) finding for a PAR (Strobel et al., 2003). Magnetic resonance imaging investigations are more expensive than diagnostic ultrasound, and may not be readily available in all areas, however when available, they may be of some value in differentiating painful ACJ conditions from other sources of shoulder pain.

Rotator Cuff Tears

Rotator cuff tears involving a large proportion of the muscle-tendon unit, or multiple tendons are of prognostic significance due to the large proportion that progress in size to the point where they may become inoperable (Gerber et al., 2011; Yamanaka & Matsumoto, 1994). Early identification enables surgery to be performed within an optimal time-frame, prior to significant tissue degeneration, that has the potential to result in an improved outcome (Gerber et al., 2011). In contrast to findings involving a PAR following diagnostic block, several individual clinical examination features were identified with the ability to both rule-in and rule-out a significant rotator cuff tear, and combinations of clinical features were also identified with high levels of accuracy for diagnosing these conditions.

Diagnostic Accuracy of the Clinical Examination for Detecting Rotator Cuff Tears

Few previous studies had investigated the diagnostic accuracy of patient history for identifying rotator cuff pathology (Chew et al., 2004; Itoi et al., 2006; Litaker et al.,

2000). Several patient history variables were found to be associated with a medium-large or multi-tendon rotator cuff tear (age >50 years, SPADI pain subscale, traumatic mechanism of pain onset, constant nature of pain and the presence of night pain). Similar to findings for the response to diagnostic blocks, the maximum sensitivity and specificity values for patient history variables (78% and 72% respectively) were lower than the corresponding maximum values for physical examination tests (96% and 97%). This supports the results of Litaker et al. (2000) in which the same history variables (history of trauma and night pain) demonstrated slightly lower maximum sensitivity (0.88) and specificity (0.73) values compared with maximum sensitivity (0.98) and specificity (0.85) values for physical examination findings (Litaker et al., 2000). However, when these history and physical examination features were combined, they contributed to the ability to both rule-in and rule-out a medium-large or multi-tendon rotator cuff tear with high levels of precision.

From the physical examination a resisted test composite (symptoms reproduced with either resisted abduction or external rotation), a passive ROM test and two orthopaedic tests were identified that were able to either rule-in or rule-out a medium-large or multi-tendon rotator cuff tear with more than 90% accuracy, indicating a combination of physical examination findings are of value for these lesions in a primary care population. Speed's test (sensitivity 0.96) and external rotation lag sign, (specificity 0.97) were also of diagnostic value for a medium-large or multi-tendon rotator cuff tear.

These results support previous findings from studies conducted in secondary care settings in which weakness or 'lag signs' (an inability to achieve or maintain a test position against gravity), were shown to be specific for loss of structural integrity of the rotator cuff (Barth et al., 2006; Castoldi et al., 2009; Hertel et al., 1996; Miller et al., 2008b; Park et al., 2005; G Walch et al., 1998). These results suggest that irrespective of health care setting, such tests are of value for identifying those patients with potentially significant rotator cuff tears. When positive these tests would provide justification for further evaluation to ascertain the extent of structural disruption, or to differentiate a neurological cause of weakness.

In contrast to previous studies, sensitivity of Speed's test in this primary care study was higher than previously reported for rotator cuff tears in secondary care populations (0.40 to 0.63) (Leroux et al., 1995; Park et al., 2005). This may be related to the high prevalence of positive tests in our study (65%) and the large number of anatomical structures placed under stress during the test. It is more commonly used to

assess long head of biceps tendon and superior labral pathology (Bennett, 1998; Holtby & Razmjou, 2004; B. W Kibler et al., 2009; Mirkovic et al., 2005), however the intimate anatomic connections between the long head of biceps, structures of the rotator interval (coracohumeral ligament, superior glenohumeral ligament, biceps pulley) and the anterior rotator cuff structures (supraspinatus and subscapularis) (Arai et al., 2010) mean that involvement of any of these structures may result in a positive test. Although only a sub-group of participants received an MR arthrogram investigation, results support the high prevalence of rotator interval pathology (55%) among those who did not report a PAR following the SAB or ACJ diagnostic block.

While two orthopaedic tests were able to rule-in or rule-out a significant rotator cuff tear, combinations of clinical features including patient history variables, resisted tests, passive ROM and orthopaedic tests demonstrated an ability to rule-in or rule-out these conditions for a higher proportion of participants. For those in whom a large rotator cuff tear is identified, referral for further evaluation of the extent of injury, or for orthopaedic consultation could be justified. When the diagnosis remains unclear, additional diagnostic imaging procedures such as diagnostic ultrasound, may be required.

In summary, individual or combinations of clinical examination features were able to both rule-in and rule-out a clinically significant rotator cuff tear in many cases. For those in whom clinical tests results are equivocal, diagnostic ultrasound imaging may be required.

Study Limitations

There were some limitations in this study. Specific limitations pertaining to individual sections of this thesis have been discussed within the relevant manuscripts including the potential for anaesthetisation of other structures during the SAB diagnostic block and the retrospective nature of some data collection relating to ultrasound findings of ACJ pathology. There are also some general limitations to consider with respect to diagnostic accuracy research that may affect interpretation of these results.

Reliability is an important aspect of test validity. Some aspects of the clinical examination require further testing to evaluate agreement between examiners for aspects of patient history and some physical examination tests before results can be widely implemented into practice. The interexaminer reliability of sonographers and radiologists for reporting of imaged pathology was not formally assessed prior to this study, however all were highly trained and experienced in musculoskeletal radiology, were from similar training backgrounds, the majority of investigations were reported by a small number of staff (n=3), and high tech equipment was used to perform all imaging procedures. Previous authors have reported high levels of agreement between examiners under similar circumstances (Naredo et al., 2006). Thus the effect of observer variability under the conditions these imaging investigations were conducted was minimised and diagnostic accuracy results with respect to imaging findings are likely to represent typical clinical practice situations.

The false-positive and false-negative response rates for diagnostic blocks around the shoulder have not previously been reported, and such investigations were beyond the scope and financial constraints of this study. However, the diagnostic blocks used as reference standard procedures in this study were performed using imaging guidance that allowed accurate placement of injections and the ability to identify those that did not reach the target structure, or in which injectate was not contained (ACJ). Thus face validity of these procedures was achieved. Assessment of the false-positive and false-negative rate of shoulder diagnostic blocks has been identified as a priority area of future research involving placebo controlled trials. Until this has been completed, results should be interpreted accordingly.

Diagnostic accuracy estimates for results for the ACJ diagnostic block may have been affected by a number of cases being excluded from the analysis in which results of the reference standard were indeterminate (pain ≤20mm on pre-injection VAS), and also

by the number of participants who reported an increase in pain following the diagnostic block. Further research is required to determine whether a substantial increase in post-injection pain intensity following an ACJ injection may itself be indicative of a symptomatic articular disorder.

Due to the limited availability of equipment and specialist radiology staff, as well as financial constraints it was not possible for all participants to receive the GHJ diagnostic block procedure in this study. This has been identified as an area for future research involving a consecutive patient series.

<u>**Clinical Applications**</u>

Results of this study have clinical practice implications for primary care clinicians involved in the assessment of painful shoulder conditions, for the diagnostic processes in primary care, and for the validity of commonly used diagnostic classification criteria.

Clinical Practice Implications

Clinical reasoning is based upon the integration of acquired skill and scientific evidence. Clinicians (often unconsciously) apply likelihood ratios to elements of clinical information to inform diagnostic reasoning and weigh the relative probability of a particular condition being present or absent. Our results provide scientific evidence that information derived from combinations of several clinical examination variables improved the ability to rule-in and rule-out subacromial structures and the ACJ as the predominant source of shoulder pain compared with individual clinical examination tests. This use of combinations of data from several clinical tests more closely approximates clinical practice, facilitating diagnostic reasoning, whereby test criteria with associated diagnostic accuracy statistics (sensitivity and specificity) can be selected that are compatible with the clinical objective. This represents integration of new scientific evidence that can be combined with the utilitarian skill of clinical reasoning to refine and improve diagnostic ability.

The specific diagnostic reasoning pathways incorporating results of this study that may be used to supplement clinical reasoning for identifying a predominant subacromial or ACJ pain source, or a clinically significant rotator cuff tear are presented in Figures 8.1 to 8.3. Further reliability testing and prospective validation is recommended before these can be widely implemented in clinical practice.

Implications for Delivery and Funding of Health Care Services:

Pending verification of results, these results have implications for multi-disciplinary health care service delivery, health-care funding policy relating to the use of diagnostic imaging and diagnostic injections, and the need for extended scope physiotherapy practitioners with the ability to perform diagnostic injection procedures and refer for diagnostic imaging procedures. The use of diagnostic injections by primary care practitioners, and more judicious use and interpretation of diagnostic imaging may result in more accurate and timely identification of those patients with painful SAB or ACJ pathology, and those who may be likely to respond to specific management

interventions. The costs of unnecessary health care services and the burden on secondary care waiting lists may thus be relieved. More efficient diagnosis and application of appropriate treatment interventions may also improve outcomes of care for primary care patients suffering from shoulder pain.

Implications for Diagnostic Classification of Shoulder Disorders

A lack of validated diagnostic criteria is a commonly reported flaw in randomised controlled trials investigating the efficacy of various interventions for the treatment of shoulder pain. Many previous trials used clinical test criteria that lack validity in the population under investigation. Clinical test criteria alone lack specificity, and imaging criteria fail to take into account the degree to which the pathology contributes to symptoms. Results of this research suggest that future diagnostic classification criteria for shoulder conditions included in intervention trials may require a combination of clinical examination and diagnostic imaging information, as well as information regarding response to diagnostic injections to identify symptomatic shoulder pathology for which different management pathways, including conservative management pathways, may exist.

Ongoing analysis will help determine whether subgroups of shoulder pain can be identified for whom specific treatment interventions may be beneficial. Such subgroups may include subgroups of SAB pathology, or combinations of SAB and rotator cuff pathology for whom targeted pain relief interventions, or combinations of injection therapy, exercise or other treatment modalities may improve outcomes.

Recommendations for Future Research

The findings from this study identified a number of areas requiring ongoing analysis or further investigation.

Ongoing Analysis

The following aspects of the study are subjects of ongoing analysis, with the intention of submitting related manuscripts for publication in due course.

- It was the intention prior to commencing this study, to also perform sub-group analyses on these results for subgroups of age, pain severity, symptom duration and psychosocial factors (fear avoidance beliefs, and mental health component scores from the SF-8 health survey). However due to the large scope of the study and analysis, this has not been conducted to date but has been prioritised for ongoing analysis and manuscript preparation.

- Evaluate the association between clinical examination and imaging predictors of response to diagnostic blocks at varying anaesthetic response levels (50%, 60%, 70%, 90% etc). Although only a small number of imaging variables were identified that were associated with an 80% PAR, it is possible that other variables may be of diagnostic and prognostic relevance using other anaesthetic response criteria.

- A number of measures of range of motion and strength (peak muscle force) were recorded in this study, however these continuous variables were not used in the analysis due to limited availability of the hand-held dynamometer and untested interexaminer reliability following methodological refinement limiting widespread clinical application. In addition selection of diagnostic cut-points for continuous data is problematic. Additional analyses involving linear regression and exploratory analysis for variable cutpoints is planned in order to determine whether range of motion or strength deficits contribute to the diagnosis of the outcomes used in this study.

- Due to the large scale of this study, the subsequent volume of data and the non-consecutive (biased) sample receiving MRA, evaluation of the diagnostic accuracy of clinical examination and imaging results for PAR to GHJ diagnostic block was not possible, however this will be conducted upon completion of this thesis.

- All participants enrolled in this study were invited to participate in a separate follow up study investigating prognostic indicators of outcome and response to corticosteroid injection. Those agreeing to participate have completed 3, 6, and 12

month follow-up visits. The 3 month follow up data is currently being analysed and papers are in an advanced stage of preparation for submission to scientific journals. The 12 month follow up data is complete and will be analysed and reported on in due course.

Future Research Recommendations

- Further studies are required to evaluate the interexaminer reliability of patient history, symptom responses and ROM and strength tests according to methods developed using the refined procedures described in Chapter 4.

- The reliability of reported x-ray and ultrasound scan results also needs to be evaluated to assess the impact of observer variability on the results of this study.

- Diagnostic injections are the best available reference standard procedure for identifying the tissue source of pain, however the construct validity of these procedures requires further evaluation using placebo controlled block procedures to determine the false-positive and false-negative response rates for SAB and ACJ diagnostic block procedures.

- It was not possible to perform the GHJ diagnostic block and magnetic resonance arthrogram procedures on all patients due to resource availability and cost constraints. To enable assessment of the diagnostic accuracy of ultrasound scan (compared with MRA findings) for rotator cuff tears, and the relationship of clinical examination, x-ray and ultrasound findings to a GHJ PAR, repeat study methodology for response to GHJ diagnostic block on consecutive series of primary care patients is required.

- As with all preliminary models, prospective validation of diagnostic test combinations is required, followed by evaluation of the accuracy of clinical and imaging models in an independent sample.

- Following validation of diagnostic clinical and imaging models, evaluation of the clinical impact of these models is required with respect to both cost-effectiveness of this process and patient outcomes compared with 'usual care'.

- Controlled trials of evidence-based primary care management strategies for shoulder pain using diagnostic classifications based upon clinical, imaging findings and diagnostic injections of local anaesthetic are required.

8.3. CONCLUSIONS

This study aimed to estimate the diagnostic accuracy of clinical examination findings for identifying sources of shoulder pain and pathology, and to assess the added diagnostic value of imaging investigations in a cohort of primary care patients. These aims were achieved, and new evidence is presented that may contribute to the ability of primary care practitioners to more accurately diagnose predominant subacromial and ACJ pain sources, and large rotator cuff tears using combinations of clinical examination findings. Ultrasound imaging features were identified that improve the ability to accurately diagnose a predominant subacromial or ACJ pain source using an 80% pain reduction standard when combined with clinical information. However, the low prevalence of these imaging findings in primary care patients with shoulder pain means that many patients may require additional diagnostic investigations such as clinically administered diagnostic injections of local anaesthetic to confirm the diagnosis.

The diagnostic process begins in primary care. Results of this study may elucidate the use of specific combinations of clinical examination tests, combined with the judicious use of diagnostic imaging and other diagnostic investigations such as diagnostic injections and imaging investigations, to accurately diagnose these disorders and enable efficient implementation of appropriate management. This approach has specific implications for the physiotherapy profession where extended scope physiotherapy practitioners are required, with the ability to refer for imaging and perform diagnostic injections. A coordinated approach to the diagnosis of shoulder pain by all medical practitioners including physiotherapists, general practitioners, radiologists, sports medicine physicians, musculoskeletal medicine doctors and orthopaedic surgeons has the potential to reduce inappropriate health care spending, reduce the burden on secondary care services and optimise the chances of a successful outcome for patients suffering from shoulder pain. The results of this study may provide a framework for a common and united approach by a range of medical practitioners to the diagnosis and management of shoulder pain in primary health care.

REFERENCES

Agre, J. C., Magness, J. L., Hull, S. Z., Wright, K. C., Baxter, T. L., Patterson, R., & Stradel, L. (1987). Strength testing with a portable dynamometer: Reliability for upper and lower extremities. *Archives of Physical Medicine & Rehabilitation, 68*, 454-458.

Akaike, H. (1974). New look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716-723.

Alasaarela, E., Takalo, R., Tervonen, O., Hakala, M., & Suramo, I. (1997). Sonography and MRI in the evaluation of painful arthritic shoulder. *British Journal of Rheumatology, 36*(9), 996-1000.

Anderson, J., Read, J. W., & Steinweg, J. (1998). The shoulder and arm. In J. Anderson, J. Steinweg, & J. W. Read (Eds.), *Atlas of Imaging in Sports Medicine* (pp. 96-129). Sydney, Australia: McGraw-Hill Book Company Australia Pty Limited.

Andren, L., & Lundberg, B. J. (1965). Treatment of rigid shoulders by joint distension during arthrography. *Acta Orthopaedica Scandinavica, 36*, 45-53.

Andrews, J. R. (2005). Diagnosis and treatment of chronic painful shoulder: Review of nonsurgical interventions. *Arthroscopy, 21*(3), 333-347. doi:10.1016/j.arthro.2004.11.003

Applegate, G. R., Hewitt, M., Snyder, S. J., Watson, E., Kwak, S., & Resnick, D. (2004). Chronic labral tears: Value of magnetic resonance arthrography in evaluating the glenoid labrum and labral-bicipital complex. *Arthroscopy, 20*(9), 959-963. doi:10.1016/j.arthro.2004.08.017

Arai, R., Mochizuki, T., Yamaguchi, K., Sugaya, H., Kobayashi, M., Nakamura, T., & Akita, K. (2010). Functional anatomy of the superior glenohumeral and coracohumeral ligaments and the subscapularis tendon in view of stabilization of the long head of the biceps tendon. *Journal of Shoulder and Elbow Surgery, 19*(1), 58-64. doi:10.1016/j.jse.2009.04.001

Ardic, F., Kahraman, Y., Kacar, M., Kahraman, M. C., Findikoglu, G., & Yorgancioglu, Z. R. (2006). Shoulder impingement syndrome: Relationships between clinical, functional and radiologic findings. *American Journal of Physical Medicine and Rehabilitation, 85*(1), 53-60. doi:10.1097/01.phm.0000179518.85484.53

Arrol, B., Robb, G., Kool, B., Astley, T., Ball, C., Carroll, M., ... Watt, J. (2004). *The diagnosis and management of soft tissue shoulder injuries and related disorders: Best practice evidence-based guideline.*: Accident Compensation Corporation.

Auge II, W. K., & Fischer, R. A. (1998). Arthroscopic distal clavicle resection for isolated atraumatic osteolysis in weight lifters. *American Journal of Sports Medicine, 26*(2), 189-192.

Auvinen, J., Tammelin, T., Taimela, S., Md, Zitting, P., & Karppinen, J. (2007). Neck and Shoulder Pains in Relation to Physical Activity and Sedentary Activities in Adolescence. *Spine, 32*(8), 1038-1044.

Awerbuch, M. S. (2008). The clinical utility of ultrasonography for rotator cuff disease, shoulder impingement syndrome and subacromial bursitis. *Medical Journal of Australia, 188*(1), 50-53.

Backhaus, M., Burmester, G. R., Gerber, T., Grassi, W., Machold, K. P., Swen, W. A., ... Manger, B. (2001). Guidelines for musculoskeletal ultrasound in rheumatology. *Annals of the Rheumatic Diseases, 60*(7), 641-649.

Balich, S. M., Sheley, R. C., Brown, T. R., Sauser, D. D., & Quinn, S. F. (1997). MR imaging of the rotator cuff tendon: Interobserver agreement and analysis of interpretive errors. *Radiology, 204*(1), 191-194.

Barber, F. A., Field, L. D., & Ryu, R. K. N. (2007). Biceps tendon and superior labrum injuries: Decision-making. *Journal of Bone & Joint Surgery (American), 89*(8), 1843-1855.

Baring, T., Emery, R., & Reilly, P. (2007). Management of rotator cuff disease: Specific treatment for specific disorders. *Best Practice and Research, 21*(2), 279-294.

Barth, J. R. H., Burkhart, S. S., & De Beer, J. F. (2006). The Bear-Hug test: A new and sensitive test for diagnosing a subscapularis tear. *Arthroscopy, 22*(10), 1076-1084. doi:10.1016/j.arthro.2006.05.005

Bartolozzi, A., Andreychik, D., & Ahmad, S. (1994). Determinants of outcome in the treatment of rotator cuff disease. *Clinical Orthopaedics and Related Research, 308*, 90-97.

Bateman, J. E. (1963). The diagnosis and treatment of ruptures of the rotator cuff. *Surgical Clinics of North America, 43*, 1523–1530.

Beaudreuil, J., Nizard, R., Thomas, T., Peyre, M., Liotard, J. P., Boileau, P., ... Walch, G. (2009). Contribution of clinical tests to the diagnosis of rotator cuff disease: A systematic literature review. *Joint Bone Spine, 76*, 15-19. doi:10.1016/j.jbspin.2008.04.015

Bencardino, J. T., Beltran, J., Rosenberg, Z. S., Rokito, A., Schmahmann, S., Mota, J., ... Rose, D. (2000). Superior labrum anterior-posterior lesions: Diagnosis with MR arthrography of the shoulder. *Radiology, 214*(1), 267-271.

Bennett, W. F. (1998). Specificity of the Speed's test: Arthroscopic technique for evaluating the biceps tendon at the level of the bicipital groove. *Arthroscopy, 14*(8), 789-796.

Binkert, C. A., Zanetti, M., Gerber, C., & Hodler, J. (2001). MR arthrography of the glenohumeral joint: Two concentrations of gadoteridol versus ringer solution as the intraarticular contrast material. *Radiology, 220*(1), 219-224.

Bisbinas, I., Belthur, M., Said, H. G., Green, M., & Learmonth, D. J. A. (2006). Accuracy of needle placement in ACJ injections. *Knee Surgery, Sports*

Traumatology, Arthroscopy, 14(8), 762-765. doi:10.1007/s00167-006-0038-5

Bjelle, A. (1989). Epidemiology of shoulder problems. *Balliere's Clinical Rheumatology, 3*(3), 437-451.

Blaine, T. A., Kim, Y. S., Voloshin, I., Chen, D., Murakami, K., Chang, S. S., ... Bigliani, L. U. (2005). The molecular pathophysiology of subacromial bursitis in rotator cuff disease. *Journal of Shoulder and Elbow Surgery, 14*, 84S-89S. doi:10.1016/j.jse.2004.09.022

Bland, J., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307-310.

Boettcher, C. E., Ginn, K. A., & Cathers, I. (2008). Standard maximum isometric voluntary contraction tests for normalizing shoulder muscle EMG. *Journal of Orthopaedic Research, 26*(12), 1591-1597. doi:10.1002/jor.20675

Bogduk, N. (2004a). Diagnostic blocks: A truth serum for malingering. *Clinical Journal of Pain, 20*(6), 409-414. doi:10.1097/00002508-200411000-00005

Bogduk, N. (2004b). *Practice guidelines for spinal diagnostic & treatment procedures* (1st ed.)[Practice Guideline]. San Francisco: International Spine Intervention Society.

Bogduk, N. (2005). Diagnosing lumbar zygapophysial joint pain. *Pain Medicine, 6*(2), 139-142. doi:10.1111/j.1526-4637.2005.05023.x

Bogduk, N. (2009). On the rational use of diagnostic blocks for spinal pain. *Neurosurgery Quarterly, 19*(2), 88-100. doi:10.1097/WNQ.0b013e3181a32e8b

Bogduk, N., Dreyfuss, P., & Govind, J. (2009). A narrative review of lumbar medial branch neurotomy for the treatment of back pain. *Pain Medicine, 10*(6), 1035-1045. doi:10.1111/j.1526-4637.2009.00692.x

Bohannon, R. W. (1999). Intertester reliability of hand-held dynamometry: A concise summary of published research. *Perceptual and Motor Skills, 88*(3), 899-902.

Boocock, M. G., Collier, J. M. K., McNair, P. J., Simmonds, M., Larmer, P. J., & Armstrong, B. (2009). A framework for the classification and diagnosis of work-related upper extremity conditions: systematic review. *Seminars in Arthritis and Rheumatism, 38*(4), 296-311. doi:10.1016/j.semarthrit.2007.10.006

Boone, D. C., Azen, S. P., Lin, C. M., Spence, C., Baron, C., & Lee, L. (1978). Reliability of goniometric measurements. *Physical Therapy, 58*, 1355-1360.

Borrero, C. G., Casagranda, B. U., Towers, J. D., & Bradley, J. P. (2010). Magnetic resonance appearance of posterosuperior labral peel back during humeral abduction and external rotation. *Skeletal Radiology, 39*(1), 19-26. doi:10.1007/s00256-009-0744-4

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., ... de Vet, H. C. W. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *British Medical Journal, 326*(7379), 41-44.

Bostrom, C., Harms-Ringdahl, K., & Nordemar, R. (1995). Relationships between measurements of impairment, disability, pain, and disease activity in rheumatoid arthritis patients with shoulder problems. *Scandinavian Journal of Rheumatology, 24*(6), 352-359.

Bot, S. D. M., van der Waal, J. M., Terwee, C. B., van der Windt, D. A. W. M., Schellevis, F. G., & Bouter, L. M. (2005). Incidence and prevalence of complaints of the neck and upper extremity in general practice. *Annals of the Rheumatic Diseases, 64*(1), 118-123. doi:10.1136/ard.2003.019349

Broadhurst, N. A. (1989). Sacroiliac dysfunction as a cause of low back pain. *Australian Family Physician, 18*(6), 623-624, 626.

Broadhurst, N. A., Gialamas, A., McElroy, H. J., & Beilby, J. J. (2004). How do Australian GPs manage shoulder dysfunction? *Australian Family Physician, 33*(10), 861-862, 864.

Bruyn, G. A. W., Naredo, E., Moller, I., Moragues, C., Garrido, J., De Bock, G. H., ... Schmidt, W. A. (2009). Reliability of ultrasonography in detecting shoulder disease in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases, 68*(3), 357-361. doi:10.1136/ard.2008.089243

Bruyn, G. A. W., Pineda, C., Hernandez-Diaz, C., Ventura-Rios, L., Moya, C., Garrido, J., ... Naredo, E. (2010). Validity of ultrasonography and measures of adult shoulder function and reliability of ultrasonography in detecting shoulder synovitis in patients with rheumatoid arthritis using magnetic resonance imaging as a gold standard. *Arthritis Care and Research, 62*(8), 1079-1086. doi:10.1002/acr.20175

Bryant, T. N. (2000). Confidence Interval Analysis [computer software]. Bristol: British Medical Journal Books. Retrieved from http://som.soton.ac.uk/cia

Buchbinder, R. (1996). Classification of soft-tissue disorders of the neck and upper limb: Do they satisfy methodological guidelines? *Journal of Clinical Epidemiology, 49*(2), 141-149.

Buchbinder, R., Green, S., & Youd, J. M. (2003). Corticosteroid injections for shoulder pain. *Cochrane Database of Systematic Reviews*(1), Art. No.: CD004016. doi:10.1002/14651858.CD004016

Burbank, K. M., Stevenson, J. H., Czarnecki, G. R., & Dorfman, J. (2008). Chronic shoulder pain: Part I. evaluation and diagnosis. *American Family Physician, 77*(4), 453-460.

Bureau, N. J., Beauchamp, M., Cardinal, E., & Brassard, P. (2006). Dynamic sonography evaluation of shoulder impingement syndrome. *American Journal of Roentgenology, 187*(1), 216-220. doi:10.2214/AJR.05.0528

Burk Jr, D. L., Karasick, D., Kurtz, A. B., Mitchell, D. G., Rifkin, M. D., Miller, C. L., ... Bartolozzi, A. R. (1989). Rotator cuff tears: Prospective comparison of MR imaging with arthrography, sonography, and surgery. *American Journal of Roentgenology, 153*(1), 87-92.

Burkhart, S. S., Barth, J. R. H., Richards, D. P., Zlatkin, M. B., & Larsen, M. (2007). Arthroscopic repair of massive rotator cuff tears with stage 3 and 4 fatty degeneration. *Arthroscopy, 23*(4), 347-354. doi:10.1016/j.arthro.2006.12.012

Burkhart, S. S., Morgan, C. D., & Kibler, B. W. (2003). The disabled throwing shoulder: Spectrum of pathology part III: The SICK scapula, scapular dyskinesis, the kinetic chain, and rehabilitation. *Arthroscopy, 19*(6), 641-661. doi:10.1016/S0749-8063(03)00389-X

Bushnell, B. D., Creighton, R. A., & Herring, M. M. (2008). The bony apprehension test for instability of the shoulder: a prospective pilot analysis. *Arthroscopy, 24*(9), 974-982. doi:10.1016/j.arthro.2008.07.019

Bussières, A. E., Peterson, C., & Taylor, J. A. M. (2008). Diagnostic imaging guideline for musculoskeletal complaints in adults-An evidence-based approach-Part 2: Upper extremity disorders. *Journal of Manipulative and Physiological Therapeutics, 31*(1), 2-32. doi:10.1016/j.jmpt.2007.11.002

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology, 46*(5), 423-429.

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Coates, M. (2011). A prospective study of shoulder pain in primary care: Prevalence of imaged pathology and response to guided diagnostic blocks. *BMC Musculoskeletal Disorders, 12*, 119. doi:10.1186/1471-2474-12-119

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Williams, M. (2011a). Reliability of a new hand-held dynamometer in measuring shoulder range of motion and strength. *Manual Therapy, 16*(1), 97-101. doi:10.1016/j.math.2010.05.005

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Taylor, S. Clinical diagnosis of a positive response to guided acromioclavicular joint diagnostic block. *(submitted)*.

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Taylor, S. Clinical predictors of a positive anaesthetic response to guided diagnostic block into the subacromial bursa. *(submitted)*

Cadogan, A., Laslett, M., Hing, W. A., McNair, P. J., & Williams, M. (2011b). Interexaminer reliability of orthopaedic special tests used in the assessment of shoulder pain. *Manual Therapy, 16*, 131-135. doi:10.1016/j.math.2010.07.009

Cahill, B. R. (1992). Atraumatic osteolysis of the distal clavicle. A review. *Sports Medicine, 13*(3), 214-222.

Calis, M., Akgun, K., Birtane, M., Karacan, I., Calis, H., & Tuzun, F. (2000). Diagnostic values of clinical diagnostic tests in subacromial impingement syndrome. *Annals of the Rheumatic Diseases, 59*(1), 44-47. doi:10.1136/ard.59.1.44

Carette, S., Marcoux, S., Truchon, R., Grondin, C., Gagnon, J., Allard, Y., & Latulippe, M. (1991). A controlled trial of corticosteroid injections into facet joints for chronic low back pain. *New England Journal of Medicine, 325*(14), 1002-1007.

Carette, S., Moffet, H., Tardif, J., Bessette, L., Morin, F., Frémont, P., ... Blanchette, C. (2003). Intraarticular corticosteroids, supervised physiotherapy, or a combination of the two in the treatment of adhesive capsulitis of the shoulder: A placebo-controlled trial. *Arthritis and Rheumatism, 48*, 829-838.

Castoldi, F., Blonna, D., & Hertel, R. (2009). External rotation lag sign revisited: Accuracy for diagnosis of full thickness supraspinatus tear. *Journal of Shoulder and Elbow Surgery, 18*(4), 529-534. doi:10.1016/j.jse.2008.11.007

Chakravarty, K. (2002). Assessment and treatment of shoulder pain. *CPD Rheumatology, 3*(2), 59-64.

Chakravarty, K., & Webley, M. (1990). Disorders of the shoulder: An often unrecognised cause of disability in elderly people. *British Medical Journal, 300*, 848-849. doi:10.1136/bmj.300.6728.848-a

Chandnani, V. P., Gagliardi, J. A., Murnane, T. G., Bradley, Y. C., DeBerardino, T. A., Spaeth, J., & Hansen, M. F. (1995). Glenohumeral ligaments and shoulder capsular mechanism: Evaluation with MR arthrography. *Radiology, 196*(1), 27-32.

Chang, C. Y., Wang, S. F., Chiou, H. J., Ma, H. L., Sun, Y. C., & Wu, H. D. (2002). Comparison of shoulder ultrasound and MR imaging in diagnosing full-thickness rotator cuff tears. *Clinical Imaging, 26*(1), 50-54.

Chang, D., Mohana-Borges, A., Borso, M., & Chung, C. B. (2008). SLAP lesions: Anatomy, clinical presentation, MR imaging diagnosis and characterization. *European Journal of Radiology, 68*(1), 72-87. doi:10.1016/j.ejrad.2008.02.026

Chard, M. D., Hazelman, R., Hazelman, B. L., King, R. H., & Reiss, B. B. (1990). Shoulder disorders in the elderly: A community survey. *Arthritis and Rheumatism, 34*(6), 766-769.

Chew, K., Pua, Y. H., Chin, J., Clarke, M., & Wong, Y. S. (2004). Clinical predictors for the diagnosis of supraspinatus pathology. *Physiotherapy Singapore, 13*(2), 12-17.

Childs, J. D., & Cleland, J. A. (2006). Development and application of clinical prediction rules to improve decision making in physical therapist practice. *Physical Therapy, 86*(1), 122-131.

Chronopoulos, E., Kim, T. K., Park, H. B., Ashenbrenner, D., & McFarland, E. G. (2004). Diagnostic value of physical tests for isolated chronic acromioclavicular lesions. *American Journal of Sports Medicine, 32*(3), 655-661. doi:10.1177/0363546503261723

Clarys, J. P., & Marfell-Jones, M. J. (1986). Anatomical segmentation in humans and the prediction of segmental masses from intra-segmental anthropometry. *Human Biology, 58*(5), 771-782.

Codman, E. A. (1934). Rupture of the supraspinatus tendon and other lesions in or about the subacromial bursa. In T. Todd (Ed.), *The Shoulder* (pp. 123-177). Boston

Cole, A., Gill, T. K., Shanahan, E. M., Phillips, P., Taylor, A. W., & Hill, C. L. (2009). Is diabetes associated with shoulder pain or stiffness? Results from a population based study. *Journal of Rheumatology, 36*(2), 371-377.

Comfort, T. H., & Arafiles, P. (1978). Barbotage of the shoulder with image-intensified fluoroscopic control of needle placement for calcific tendinitis. *Clinical Orthopaedics and Related Research, 135*, 171-178.

Connell, D. A., Potter, H. G., Wickiewicz, T. L., Altchek, D. W., & Warren, R. F. (1999). Noncontrast magnetic resonance imaging of superior labral lesions. 102 cases confirmed at arthroscopic surgery. *American Journal of Sports Medicine, 27*(2), 208-213.

Connor, P. M., Banks, D. M., Tyson, A. B., Coumas, J. S., & Alessandro, D. F. (2003). Magnetic resonance imaging of the asymptomatic shoulder of overhead athletes. *American Journal of Sports Medicine, 31*(5), 724-727.

Constant, C. R., & Murley, A. H. G. (1987). A clinical method of functional assessment of the shoulder. *Clinical Orthopaedics and Related Research, 214*, 160-164.

Croft, P. R., Pope, D. P., & Silman, A. J. (1996). The clinical course of shoulder pain: Prospective cohort study in primary care. *British Medical Journal, 313*(7057), 601-602.

Cullen, D. M., Breidahl, W. H., & Janes, G. C. (2007). Diagnostic accuracy of shoulder ultrasound performed by a single operator. *Australasian Radiology, 51*(3), 226-229. doi:10.1111/j.1440-1673.2007.01685.x

Cummins, C. A., Sasso, L. M., & Nicholson, D. (2009). Impingement syndrome: Temporal outcomes of nonoperative treatment. *Journal of Shoulder and Elbow Surgery, 18*(2), 172-177. doi:10.1016/j.jse.2008.09.005

Cvitanic, O., Tirman, P. F. J., Feller, J. F., Bost, F. W., Minter, J., & Carroll, K. W. (1997). Using abduction and external rotation of the shoulder to increase the sensitivity of MR arthrography in revealing tears of the anterior glenoid labrum. *American Journal of Roentgenology, 169*(3), 837-844.

Cyriax, J. H. (1978). *Textbook of Orthopaedic Medicine: Diagnosis of Soft Tissue Lesions* (7th ed., Vol. One). London: Bailliere Tindall.

Cyriax, J. H. (1982). *Textbook of Orthopaedic Medicine, Volume 1: Diagnosis of Soft Tissue Lesions* (8th ed.). Minneapolis: Balliere Tindall.

Datta, S., Lee, M., Falco, F. J., Bryce, D. A., & Hayek, S. M. (2009). Systematic assessment of diagnostic accuracy and therapeutic utility of lumbar facet joint interventions. *Pain Physician, 12*(2), 437-460.

Davidson, M. (2002). The interpretation of diagnostic tests: A primer for physiotherapists. *Australian Journal of Physiotherapy, 48*, 227-233.

Davis, T. R. C. (1998). Diagnostic criteria for upper limb disorders in epidemiological studies. *Journal of Hand Surgery (British and European), 23*(5), 567-569.

De Conti, G., Marchioro, U., Dorigo, A., Boscolo, N., Vio, S., Trevisan, M., ... Angelini, F. (2010). Percutaneous ultrasound-guided treatment of shoulder tendon calcifications: Clinical and radiological follow-up at 6 months. *Journal of Ultrasound, 13*(4), 188-198. doi:10.1016/j.jus.2010.10.012

de Jesus, J. O., Parker, L., Frangos, A. J., Nazarian, L. N., de Jesus, J. O., Parker, L., ... Nazarian, L. N. (2009). Accuracy of MRI, MR arthrography, and ultrasound in the diagnosis of rotator cuff tears: A meta-analysis. *American Journal of Roentgenology, 192*(6), 1701-1707. doi:10.2214/AJR.08.1241

De Maeseneer, M., Van Roy, F., Lenchik, L., Shahabpour, M., Jacobson, J., Ryu, K. N., ... Osteaux, M. (2000). CT and MR arthrography of the normal and pathologic anterosuperior labrum and labral-bicipital complex. *Radiographics, 20 (Suppl)*, 67-81.

de Winter, A. F., Jans, M. P., Scholten, J. P., Deville, W., van Schaardenburg, D., & Bouter, L. M. (1999). Diagnostic classification of shoulder disorders: Interobserver agreement and determinants of disagreement. *Annals of the Rheumatic Diseases, 58*, 272-277.

Deitch, J. R. (2004). Acromioclavicular joint injuries. *Current Opinion in Orthopaedics, 15*(4), 261-266. doi:10.1097/00001433-200408000-00012

Dessaur, W. A., & Magarey, M. E. (2008). Diagnostic accuracy of clinical tests for superior labral anterior posterior lesions: A systematic review. *Journal of Orthopaedic & Sports Physical Therapy, 38*(6), 341-352. doi:10.2519/jospt.2008.2768

Diehr, S., Ison, D., & Jamieson, B. (2006). What is the best way to diagnose a suspected rotator cuff tear? *Journal of Family Practice, 55*(7), 621-624.

Dinauer, P. A., Flemming, D. J., Murphy, K. P., & Doukas, W. C. (2007). Diagnosis of superior labral lesions: Comparison of noncontrast MRI with indirect MR arthrography in unexercised shoulders. *Skeletal Radiology, 36*(3), 195-202. doi:10.1007/s00256-006-0237-7

Dinnes, J., Loveman, E., McIntyre, L., & Waugh, N. (2003). The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: A systematic review. *Health Technology Assessment, 7*(29), 1-178.

Donohoe, M. T., Kravitz, R. L., Wheeler, D. B., Chandra, R., Chen, A., & Humphries, N. (1999). Reasons for outpatient referrals from generalists to specialists. *Journal of General Internal Medicine, 14*(5), 281-286. doi:10.1046/j.1525-1497.1999.00324.x

Dreyfuss, P., Halbrook, B., Pauza, K., Joshi, A., McLarty, J., & Bogduk, N. (2000). Efficacy and validity of radiofrequency neurotomy for chronic lumbar zygapophysial joint pain. *Spine, 25*(10), 1270-1277. doi:10.1097/00007632-200005150-00012

Dreyfuss, P., Michaelsen, M., Pauza, K., McLarty, J., & Bogduk, N. (1996). The value of history and physical examination in diagnosing sacroiliac joint pain. *Spine 21*, 2594–2602.

Ebell, M. H. (2005). Diagnosing rotator cuff tears. *American Family Physician, 71*(8), 1587-1588.

Eberhardt, K. B., & Fex, E. (1995). Functional impairment and disability in early rheumatoid arthritis- Development over 5 years. *Journal of Rheumatology, 22*(6), 1037-1042.

Ellenbecker, T. S. (1996). Muscular strength relationship between normal grade manual muscle testing and isokinetic measurement of shoulder internal and external rotators. *Isokinetics and Exercise Science, 6*, 51-56.

Falco, F. J. E., Erhart, S., Wargo, B. W., Bryce, D. A., Atluri, S., Datta, S., & Hayek, S. M. (2009). Systematic review of diagnostic utility and therapeutic effectiveness of cervical facet joint interventions. *Pain Physician, 12*(2), 323-344.

Farin, P. U., Jaroma, H., Harju, A., & Soimakallio, S. (1990). Shoulder impingement syndrome: Sonographic evaluation. *Radiology, 176* (3), 845-849.

Farley, T. E., Neumann, C. H., Steinbach, L. S., Jahnke, A. J., & Petersen, S. S. (1992). Full-thickness tears of the rotator cuff of the shoulder: Diagnosis with MR imaging. *American Journal of Roentgenology, 158*(2), 347-351.

Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I: the problems of two paradoxes. *Journal of Clinical Epidemiology, 43*(6), 543-549.

Feleus, A., Bierma-Zeinstra, S. M., Miedema, H. S., Bernsen, R. M., Berhaar, J. A., & Koes, B. W. (2008). Incidence of non-traumatic complaints of arm, neck and shoulder in general practice. *Manual Therapy, 13*, 426-433. doi:10.1016/j.math.2007.05.010

Flahault, A., Cadilhac, M., & Thomas, G. (2005). Sample size calculation should be performed for design accuracy in diagnostic test studies. *Journal of Clinical Epidemiology, 58*(8), 859-862. doi:10.1016/j.jclinepi.2004.12.009

Fotiadou, A. N., Vlychou, M., Papadopoulos, P., Karataglis, D. S., Palladas, P., & Fezoulidis, I. V. (2008). Ultrasonography of symptomatic rotator cuff tears

compared with MR imaging and surgery. *European Journal of Radiology, 68*(1), 174-179. doi:10.1016/j.ejrad.2007.11.002

Frei, R., Chladek, P., Trc, T., Kopecny, Z., & Kautzner, J. (2008). Arthroscopic evaluation of ultrasonography and magnetic resonance imaging for diagnosis of rotator cuff tear. *Ortopedia Traumatologia Rehabilitacja, 10*(2), 111-114.

Fritz, J. M., & Wainner, R. S. (2001). Examining diagnostic tests: An evidence-based perspective. *Physical Therapy, 81*(9), 1546-1564.

Funk, L., & Snow, M. (2007). SLAP tears of the glenoid labrum in contact athletes. *Clinical Journal of Sport Medicine, 17*(1), 1-4.

Funke, M., Kopka, L., Vosshenrich, R., Oestmann, J. W., & Grabbe, E. (1996). MR arthrography in the diagnosis of rotator cuff tears: Standard spin-echo alone or with fat suppression? *Acta Radiologica, 37*(5), 627-632.

Gajdosik, R. L., & Bohannon, R. W. (1987). Clinical measurement of range of motion; review of goniometry emphasizing reliability and validity. *Physical Therapy, 67*, 1867-1872.

Garcilazo, C., Cavallasca, J. A., & Musuruana, J. L. (2010). Shoulder manifestations of diabetes mellitus *Current Diabetes Reviews, 6*(5), 334-340.

Gaskill, T. R., Braun, S., & Millett, P. J. (2010). The rotator interval: Pathology and management. *Arthroscopy, 27*(4), 556-567. doi:10.1016/j.arthro.2010.10.004

Gerber, C., Fuchs, B., & Hodler, J. (2000). The results of repair of massive tears of the rotator cuff. *Journal of Bone and Joint Surgery (American), 82*(4), 505-515.

Gerber, C., Galantay, R. V., & Hersche, O. (1998). The pattern of pain produced by irritation of the acromioclavicular joint and the subacromial space. *Journal of Shoulder and Elbow Surgery, 7*(4), 352-355.

Gerber, C., Hersche, O., & Farron, A. (1996). Isolated rupture of the subscapularis tendon: results of operative repair. *Journal of Bone and Joint Surgery (American), 78A*(7), 1015-1023.

Gerber, C., Wirth, S. H., & Farshad, M. (2011). Treatment options for massive rotator cuff tears. *Journal of Shoulder and Elbow Surgery, 20*(2), S20-S29. doi:10.1016/j.jse.2010.11.028

Gill, H. S., El Rassi, G., Bahk, M. S., Castillo, R. C., & McFarland, E. G. (2007). Physical examination for partial tears of the biceps tendon. *American Journal of Sports Medicine, 35*(8), 1334-1340. doi:10.1177/0363546507300058

Ginn, K. A., Cohen, M. L., & Herbert, R. D. (2006). Does hand-behind-back range of motion accurately reflect shoulder internal rotation? *Journal of Shoulder and Elbow Surgery, 15*(3), 311-314. doi:10.1016/j.jse.2005.08.005

Goldberg, J. A., Bruce, W. J. M., Walsh, W., & Sonnabend, D. H. (2003). Role of community diagnostic ultrasound examination in the diagnosis of full-thickness rotator cuff tears. *ANZ Journal of Surgery, 73*(10), 797-799. doi:10.1046/j.1445-2197.2003.02790.x

Gotoh, M., Hamada, K., Yamakawa, H., & al, e. (2001). Interleukin-I-induced subacromial synovitis and shoulder pain in rotator cuff diseases. *Rheumtology, 40*(9), 995-1001.

Green, R., Shanley, K., Taylor, N. F., & Perrott, M. (2008). The anatomical basis for clinical tests assessing musculoskeletal function of the shoulder. *Physical Therapy Reviews, 13*(1), 17-24. doi:10.1179/174328808X251966

Green, S., Buchbinder, R., & Hetrick, S. (2007). Physiotherapy interventions for shoulder pain. *Cochrane Database of Systematic Reviews, 2*.

Green, S., Buchbinder, R., & Hetrick, S. E. (2003). Physiotherapy interventions for shoulder pain. *Cochrane Database of Systematic Reviews*(2), Art. No.: CD004258. doi:10.1002/14651858.CD004258

Guanche, C. A., & Jones, D. C. (2003). Clinical testing for tears of the glenoid labrum. *Arthroscopy, 19*(5), 517-523. doi:10.1053/jars.2003.50104

Gusmer, P. B., Potter, H. G., Schatz, J. A., Wickiewicz, T. L., Altchek, D. W., O'Brien, S. J., & Warren, R. F. (1996). Labral injuries: Accuracy of detection with unenhanced MR imaging of the shoulder. *Radiology, 200*(2), 519-524.

Haddock, M.-C., & Funk, L. (2006). Labral tears in rock climbers. *Clinical Journal of Sport Medicine, 16*(3), 271-273.

Hagberg, M., & Wegman, D. H. (1987). Prevalence rates and odds ratios of shoulder-neck diseases in different occupational groups. *British Journal of Industrial Medicine, 44*(9), 602-610.

Hammar, M. V., Wintzell, G. B., Gunnar, K., Åström, O., Larsson, S., & Elvin, A. (2001). Role of US in the preoperative evaluation of patients with anterior shoulder instability. *Radiology, 219*(1), 29-34.

Harrington, J. M., Carter, J. T., Birrell, L., & Gompertz, D. (1998). Surveillance case definitions for work related upper limb pain syndromes. *Occupational and Environmental Medicine, 55*, 264-271.

Hawkins, R., & Kennedy, J. (1980). Impingement syndrome in athletes. *American Journal of Sports Medicine, 8*, 151-158.

Hayes, K., & Petersen, C. M. (2001). Reliability of assessing end-feel and pain and resistance sequence in subjects with painful shoulders and knees. *Journal of Orthopaedic & Sports Physical Therapy, 31*(8), 432-445.

Hayes, K., Walton, J., Szomor, Z. L., & Murrell, G. A. C. (2002). Reliability of 3 methods for assessing shoulder strength. *Journal of Shoulder and Elbow Surgery, 11*(1), 33-39.

Hayes, K., Walton, J. R., Szomor, Z. R., & Murrell, G. A. (2001). Reliability of five methods for assessing shoulder range of motion. *Australian Journal of Physiotherapy, 47*(4), 289-294.

Hayes, K. W., & Petersen, C. M. (2003). Reliability of classifications derived from Cyriax's resisted testing in subjects with painful shoulders and knees. *Journal of Orthopaedic and Sports Physical Therapy, 33*(5), 235-246.

Hegedus, E. J., Goode, A., Campbell, S., Morin, A., Tamaddoni, M., Moorman, C. T., & Cook, C. (2008). Physical examination tests of the shoulder: A systematic review with meta-analysis of individual tests. *British Journal of Sports Medicine*(42), 80-92. doi:10.1136/bjsm.2007.038406

Helliwell, P. S., Bennett, R. M., Littlejohn, G., Muirden, K. D., & Wigley, R. D. (2003). Towards epidemiological criteria for soft-tissue disorders of the arm. *Occupational Medicine, 53*(5), 313-319.

Henkus, H. E., Cobben, L. P. J., Coerkamp, E. G., Nelissen, R. G. H. H., & Van Arkel, E. R. A. (2006). The accuracy of subacromial injections: A prospective randomized magnetic resonance imaging study. *Arthroscopy, 22*(3), 277-282. doi:10.1016/j.arthro.2005.12.019

Hertel, R., Ballmer, F. T., Lambert, S. M., & Berber, C. (1996). Lag signs in the diagnosis of rotator cuff rupture. *Journal of Shoulder and Elbow Surgery, 5*(4), 307-313.

Heyworth, B. E., & Williams, R. J. I. I. I. (2009). Internal impingement of the shoulder. *American Journal of Sports Medicine, 37*(5), 1024-1037. doi:10.1177/0363546508324966

Hirano, Y., Sashi, R., Izumi, J., Itoi, E., & Watarai, J. (2006). Comparison of the MR findings on indirect MR arthrography in patients with rotator cuff tears with and without symptoms. *Radiation Medicine, 24*(1), 23-27. doi:10.1007/BF02489985

Hodler, J., Kursunoglu-Brahme, S., Snyder, S. J., Cervilla, V., Karzel, R. P., Schweitzer, M. E., ... Resnick, D. (1992). Rotator cuff disease: Assessment with MR arthrography versus standard MR imaging in 36 patients with arthroscopic confirmation. *Radiology, 182*(2), 431-436.

Holtby, R., & Razmjou, H. (2004). Accuracy of the Speed's and Yergason's tests in detecting biceps pathology and SLAP lesions: Comparison with arthroscopic findings. *Arthroscopy, 20*(3), 231-236.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). New York: Wiley. doi:10.1002/0471722146

Hoving, J. L., Buchbinder, R., Green, S., Forbes, A., Bellamy, N., Brand, C., ... Stockman, A. (2002). How reliable do rheumatologists measure shoulder movement? *Annals of Rheumatic Diseases, 61*, 612-616.

Hughes, P. C., Taylor, N. F., & Green, R. A. (2008). Most clinical tests cannot accurately diagnose rotator cuff pathology: A systematic review. *Australian Journal of Physiotherapy, 54*, 159-170.

Iagnocco, A., Coari, G., Leone, A., & Valesini, G. (2003). Sonographic study of painful shoulder. *Clinical and Experimental Rheumatology, 21*(3), 355-358.

Iagnocco, A., Filippucci, E., Meenagh, G., Delle Sedie, A., Riente, L., Bombardieri, S., ... Valesini, G. (2006). Ultrasound imaging for the rheumatologist I: Ultrasonography of the shoulder. *Clinical and Experimental Rheumatology, 24*(1), 6-11.

Iannotti, J. P., Ciccone, J., Buss, D. D., Visotsky, J. L., Mascha, E., Cotman, K., & Rawool, N. M. (2005). Accuracy of office-based ultrasonography of the shoulder for the diagnosis of rotator cuff tears. *Journal of Bone and Joint Surgery (American), 87*(6), 1305-1311. doi:10.2106/jbjs.d.02100

Iannotti, J. P., Zlatkin, M. B., Esterhai, J. L., Kressel, H. Y., Dalinka, M. K., & Spindler, K. P. (1991). Magnetic resonance imaging of the shoulder. Sensitivity, specificity, and predictive value. *Journal of Bone and Joint Surgery (American), 73*(1), 17-29.

Ide, K., Shirai, Y., & Ito, H. (1996). Sensory nerve supply in the human subacromial bursa. *Journal of Shoulder and Elbow Surgery, 5*(5), 371-382.

Information Services ACC. (2011). *Accident compensation corporation shoulder injury costs*. Wellington, New Zealand: Accident Compensation Corporation [ACC].

International Spine Intervention Society [ISIS]. (2004). *Practice guidelines for spinal diagnostic & treatment procedures* San Francisco.

Itoi, E., Kido, T., Sano, A., Urayama, M., & Sato, K. (1999). Which is more useful, the full can test or the empty can test in detecting the torn supraspinatus tendon? *American Journal of Sports Medicine, 27*(1), 65-68.

Itoi, E., Minagawa, H., Sato, T., Sato, K., & Tabata, S. (1997). Isokinetic strength after tears of the supraspinatus tendon. *Journal of Bone and Joint Surgery (British), 79*(1), 77-82. doi:10.1302/0301-620x.79b1.6860

Itoi, E., Minagawa, H., Yamamoto, N., Seki, N., & Abe, H. (2006). Are pain location and physical examinations useful in locating a tear site of the rotator cuff? *American Journal of Sports Medicine, 34*(2), 256-264. doi:10.1177/0363546505280430

Jaeschke, R., Guyatt, G. H., & Sackett, D. L. (1994). User's 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? *Journal of the American Medical Association, 271*(9), 703-707.

Jee, W. H., McCauley, T. R., Katz, L. D., Matheny, J. M., Ruwe, P. A., & Daigneault, J. P. (2001). Superior labral anterior posterior (SLAP) lesions of the glenoid labrum: Reliability and accuracy of MR arthrography for diagnosis. *Radiology, 218*(1), 127-132.

Jia, X., Ji, J.-H., Petersen, S. A., Keefer, J., & McFarland, E. G. (2008). Clinical evaluation of the shoulder shrug sign. *Clinical Orthopaedics & Related Research, 466*(11), 2813-2819. doi:10.1007/s11999-008-0331-3

Jin, W., Ryu, K. N., Kwon, S. H., Rhee, Y. G., & Yang, D. M. (2006). MR arthrography in the differential diagnosis of type II superior labral anteroposterior lesion and sublabral recess. *American Journal of Roentgenology, 187*(4), 887-893. doi:10.2214/ajr.05.0955

Jin, W., Ryu, K. N., Kwon, S. H., Rhee, Y. G., & Yang, D. M. (2006). MR arthrography in the differential diagnosis of Type II superior labral anteroposterior lesion and sublabral recess. *American Journal of Roentgenology, 187*, 887-893.

Jobe, F., & Moynes, D. (1982). Delineation of diagnostic criteria and a rehabilitation program for rotator cuff injuries. *American Journal of Sports Medicine, 10*, 336-339.

Jobe, F. W., & Kvitne, R. S. (1989). Shoulder pain in the overhand or throwing athlete: The relationship of anterior instability and rotator cuff impingement. *Orthopaedic Review, 18*(9), 963-975.

Johal, P., Martin, D., Broadhurst, N., Johal, P., Martin, D., & Broadhurst, N. (2008). Managing shoulder pain in general practice - assessment, imaging and referral. *Australian Family Physician, 37*(4), 263-265.

Johansen, J. A., Grutter, P. W., McFarland, E. G., & Petersen, S. A. (2011). Acromioclavicular joint injuries: Indications for treatment and treatment options. *Journal of Shoulder and Elbow Surgery, 20*(2), S70-S82. doi:10.1016/j.jse.2010.10.030

Johansson, K., & Ivarson, S. (2009). Intra- and interexaminer reliability of four manual shoulder maneuvers used to identify subacromial pain. *Manual Therapy, 14*(2), 231-239. doi:10.1016/j.math.2008.03.003

Jones, G. L. M. D., & Galluch, D. B. M. D. (2007). Clinical assessment of superior glenoid labral lesions: A systematic review. *Clinical Orthopaedics & Related Research, 455*, 45-51.

Joo, H. O., Jae, Y. K., Woo, S. K., Hyun, S. G., & Ji, H. L. (2008). The evaluation of various physical examinations for the diagnosis of type II superior labrum anterior and posterior lesion. *American Journal of Sports Medicine, 36*(2), 353-359. doi:10.1177/0363546507308363

Kang, C. H., Kim, S. S., Kim, J. H., Chung, K. B., Kim, Y. H., Oh, Y.-W., ... Kim, B. H. (2009). Supraspinatus tendon tears: Comparison of 3D US and MR arthrography with surgical correlation. *Skeletal Radiology, 38*(11), 1063-1069. doi:10.1007/s00256-009-0734-6

Kaplan, P. A., & Resnick, D. (1986). Stress-induced osteolysis of the clavicle. *Radiology, 158*(1), 139-140.

Kayser, R., Hampf, S., Pankow, M., Seeber, E., & Heyde, C. E. (2005). Validity of ultrasound examinations of disorders of the shoulder joint. *Ultraschall in der Medizin, 26*(4), 291-298. doi:10.1055/s-2005-858525

Keener, J. D., Wei, A. S., Kim, H. M., Steger-May, K., & Yamaguchi, K. (2009). Proximal humeral migration in shoulders with symptomatic and

asymptomatic rotator cuff tears. *Journal of Bone and Joint Surgery (American), 91*(6), 1405-1413. doi:10.2106/JBJS.H.00854

Kelly, A. M., & Fessell, D. (2009). Ultrasound compared with magnetic resonance imaging for the diagnosis of rotator cuff tears: A critically appraised topic. *Seminars in Roentgenology, 44*(3), 196-200.

Kessel, L., & Watson, M. (1977). The painful arc syndrome: Clinical classification as a guide to management. *Journal of Bone and Joint Surgery (British), 59*, 166-172.

Kibler, B. W. (1998). The role of the scapula in athletic shoulder function. *American Journal of Sports Medicine, 26*(2), 325-337.

Kibler, B. W., Sciascia, A. D., Hester, P., Dome, D., & Jacobs, C. (2009). Clinical utility of traditional and new tests in the diagnosis of biceps tendon injuries and superior labrum anterior and posterior lesions in the shoulder. *The American Journal of Sports Medicine, 37*(9), 1840-1847. doi:10.1177/0363546509332505

Kim, S., Ha, K., & Han, K. (1999). Biceps load test: A clinical test for superior labrum anterior and posterior lesions in shoulders with recurrent anterior dislocations. *American Journal of Sports Medicine, 27*(3), 300-303.

Kim, S. H., Ha, K. I., Ahn, J. H., Kim, S. H., & Choi, H. J. (2001). Biceps load test II: A clinical test for SLAP lesions of the shoulder. *Arthroscopy, 17*(2), 160-164.

Kim, S. H., Park, J. C., Jeong, W. K., & Shin, S. K. (2005). The Kim test: A novel test for posteroinferior labral lesion of the shoulder - a comparison to the Jerk test. *American Journal of Sports Medicine, 33*(8), 1188-1192. doi:10.1177/0363546504272687

Kim, S. H., Park, J. C., & Oh, I. (2004). Painful Jerk test: A predictor of success in nonoperative treatment of posteroinferior instability of the shoulder. *American Journal of Sports Medicine, 32*(8), 1849-1855. doi:10.1177/0363546504265263

Kim, Y. S., Kim, J. M., Ha, K. Y., Choy, S., Joo, M. W., & Chung, Y. G. (2007). The passive compression test: A new clinical test for superior labral tears of the shoulder. *American Journal of Sports Medicine, 35*(9), 1489-1494. doi:10.1177/0363546507301884

King, L. J., & Healy, J. (1999). Imaging of the painful shoulder. *Manual Therapy, 4*(1), 11-18.

Kolber, M. J., Beekhuizen, K., Cheng, M.-S. S., & Fiebert, I. M. (2007). The reliability of hand-held dynamometry in measuring isometric strength of the shoulder internal and external rotator musculature using a stabilisation device. *Physiotherapy Theory & Practice, 23*(2), 119-124. doi:10.1080/09593980701213032

Kolla, S., & Motamedi, K. (2007). Ultrasound evaluation of the shoulder. *Seminars in Musculoskeletal Radiology, 11*(2), 117-125. doi:10.1055/s-2007-1001877

Krief, O. P., & Huguet, D. (2006). Shoulder pain and disability: Comparison with MR findings. *American Journal of Roentgenology, 186*, 1234-1239. doi:10.2214/AJR.04.1766

Kuhn, M. A., & McGuigan, F. (2006). Magnetic resonance imaging assessment of the accuracy of subacromial injections. *Arthroscopy, 22*(6, Supplement 1), e16. doi:10.1016/j.arthro.2006.04.031

Kuijpers, T., van der Windt, D. A. W. M., Boeke, A. J. P., Twisk, J. W. R., Vergouwe, Y., Bouter, L. M., & van der Heijden, G. J. M. G. (2006). Clinical prediction rules for the prognosis of shoulder pain in general practice. *Pain, 120*(3), 276-285. doi:10.1016/j.pain.2005.11.004

Lähteenmäki, H. E., Hiltunen, A., Virolainen, P., & Nelimarkka, O. (2007). Repair of full-thickness rotator cuff tears is recommended regardless of tear size and age: A retrospective study of 218 patients. *Journal of Shoulder and Elbow Surgery, 16*(5), 586-590. doi:10.1016/j.jse.2007.04.001

Landis, J. R., & Koch, G. C. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Lange, U., Piegsa, M., Teichmann, J., & Neeck, G. (2000). Ultrasonography of the glenohumeral joints: A helpful instrument in differentiation in elderly onset rheumatoid arthritis and polymyalgia rheumatica. *Rheumatology International, 19*(5), 185-189.

Larson, H. M., O'Connor, F. G., & Nirschl, R. P. (1996). Shoulder pain: The role of diagnostic injections. *American Family Physician, 53*(5), 1637-1643.

Laslett, M. (1996). The Shoulder Complex. In M. Laslett (Ed.), *Mechanical Diagnosis and Therapy. The Upper Limb* (pp. 99-158): Library of Congress Cataloging in Publication Data.

Laslett, M. (2006). Pain provocation tests for diagnosis of sacroiliac joint pain. *Australian Journal of Physiotherapy, 52*(3), 229.

Laslett, M., Aprill, C. N., McDonald, B., & Oberg, B. (2006). Clinical predictors of lumbar provocation discography: A study of clinical predictors of lumbar provocation discography. *European Spine Journal, 15*(10), 1473-1484. doi:doi: 10.1007/s00586-006-0062-7

Laslett, M., Aprill, C. N., McDonald, B., & Young, S. (2005). Diagnosis of sacroiliac joint pain: validity of individual provocation tests and composites of tests. *Manual Therapy, 10*, 207-218. doi:10.1016/j.math.2005.01.003

Laslett, M., McDonald, B., Aprill, C. N., Tropp, H., & Oberg, B. (2006). Clinical predictors of screening lumbar zygapophyseal joint blocks: development of clinical prediction rules. *Spine Journal, 6*(4), 370-379. doi:10.1016/j.spinee.2006.01.004

Laupacis, A., Sekar, N., & Stiell, I. G. (1997). Clinical prediction rules: A review and suggested modifications of methodological standards. *Journal of the American Medical Association, 277*(6), 488-494.

Le Corroller, T., Cohen, M., Aswad, R., Pauly, V., & Champsaur, P. (2008). Sonography of the painful shoulder: Role of the operator's experience. *Skeletal Radiology, 37*(11), 979-986. doi:10.1007/s00256-008-0539-z

Lea, R., & Gerhardt, J. (1995). Range of motion measurements. *Journal of Bone & Joint Surgery (American), 77-A*(5), 784-798.

Leeflang, M. M. G., Bossuyt, P. M. M., & Irwig, L. (2009). Diagnostic test accuracy may vary with prevalence: Implications for evidence-based diagnosis. *Journal of Clinical Epidemiology, 62*(1), 5-12. doi:10.1016/j.jclinepi.2008.04.007

Legan, J. M., Burkhard, T. K., Goff Ii, W. B., Balsara, Z. N., Martinez, A. J., Burks, D. D., ... Lapoint, J. M. (1991). Tears of the glenoid labrum: MR imaging of 88 arthroscopically confirmed cases. *Radiology, 179*(1), 241-246.

Leggin, B. G., Neuman, R. M., Iannotti, J. P., Williams, G. R., & Thompson, E. C. (1996). Intrarater and interrater reliability of three isometric dynamometers in assessing shoulder strength. *Journal of Shoulder and Elbow Surgery, 5*(1), 18-24.

Lentz, T. A., Barabas, J. A., Day, T., Bishop, M. D., & George, S. Z. (2009). The relationship of pain intensity, physical impairment, and pain-related fear to function in patients with shoulder pathology. *Journal of Orthopaedic & Sports Physical Therapy, 39*(4), 270-277.

Leroux, J., Thomas, E., Bonnel, F., & Blotman, F. (1995). Diagnostic value of clinical tests for shoulder impingement syndrome. *Revue du Rhumatisme (Engl Ed), 62*, 423-428.

Levy, H. J., Gardner, R. D., & Lemak, L. J. (1991). Arthroscopic subacromial decompression in the treatment of full-thickness rotator cuff tears. *Arthroscopy, 7*(1), 8-13. doi:10.1016/0749-8063(91)90071-5

Lijmer, J. G., Mol, B. W., Heisterkamp, S., Bonsel, G. J., Prins, M. H., Van Der Meulen, J. H. P., & Bossuyt, P. M. M. (1999). Empirical evidence of design-related bias in studies of diagnostic tests. *Journal of the American Medical Association, 282*(11), 1061-1066. doi:10.1001/jama.282.11.1061

Lin, J., Hanten, W. P., Olson, S. L., Roddey, T. S., Soto-quijano, D. A., Lim, H. K., & Sherwood, A. M. (2005). Functional activity characteristics of individuals with shoulder dysfunctions. *Journal of Electromyography & Kinesiology, 15*(6), 576-586.

Lindgren, I., Jonsson, A.-C., Norrving, B., & Lindgren, A. (2007). Shoulder pain after stroke: a prospective population-based study. *Stroke, 38*(2), 343-348.

Linsell, L., Dawson, J., Zondervan, K., Rose, P., Randall, T., Fitzpatrick, R., & Carr, A. (2006). Prevalence and incidence of adults consulting for shoulder conditions in UK primary care: Patterns of diagnosis and referral. *Rheumatology, 45*(2), 215-221. doi:10.1093/rheumatology/kei139

Litaker, D., Pioro, M., El Bilbeisi, H., & Brems, J. (2000). Returning to the bedside: Using the history and physical examination to identify rotator cuff tears. *Journal of the American Geriatric Society, 48*(12), 1633-1637.

Liu, S. H., Henry, M. H., Nuccion, S., S, S. M., & Dorey, F. (1996). Diagnosis of glenoid labral tears- A comparison between magnetic resonance imaging and clinical examinations. *American Journal of Sports Medicine, 24*(2), 149-154.

Lo, I. K. Y., Nonweiler, B., Woolfrey, M., Litchfield, R., & Kirkley, A. (2004). An evaluation of the apprehension, relocation, and surprise tests for anterior shoulder instability. *American Journal of Sports Medicine, 32*(2), 301-307. doi:10.1177/0095399703258690

Loew, R., Kreitner, K. F., Runkel, M., Zoellner, J., & Thelen, M. (2000). MR arthrography of the shoulder: Comparison of low-field (0.2 T) vs high-field (1.5 T) imaging. *European Radiology, 10*(6), 989-996.

Luime, J. J., Koes, B. W., Hendriksen, I. J. M., Burdorf, A., Verhagen, A. P., Miedema, H. S., & Verhaar, J. A. (2004). Prevalence and incidence of shoulder pain in the general population: A systematic review. *Scandinavian Journal of Rheumatology, 33*(2), 73-81. doi:10.1080/03009740310004667

Luime, J. J., Verhagen, A. P., Miedema, H. S., Kuiper, J. I., Burdorf, A., Verhaar, J. A., & Koes, B. W. (2004). Does this patient have an instability of the shoulder or a labrum lesion? *Journal of the American Medical Association, 292*(16), 1989-1999.

MacDermid, J. C., Chesworth, B. M., Patterson, S., & Roth, J. H. (1999). Intratester and intertester reliability of goniometric measurement of passive lateral shoulder rotation. *Journal of Hand Therapy, 12*, 187-192.

MacDermid, J. C., Ghobrial, M., Quirion, K. B., St-Amour, M., Tsui, T., Humphreys, D., ... Galea, V. (2007). Validation of a new test that assesses functional performance of the upper extremity and neck (FIT-HaNSA) in patients with shoulder pathology. *BMC Musculoskeletal Disorders, 8*, 42. doi:10.1186/1471-2474-8-42

MacDermid, J. C., Ramos, J., Drosdowech, D., Faber, K., & Patterson, S. (2004). The impact of rotator cuff pathology on isometric and isokinetic strength, function, and quality of life. *Journal of Shoulder & Elbow Surgery, 13*(6), 593-598.

MacDermid, J. C., Solomon, P., & Prkachin, K. (2006). The Shoulder Pain and Disability Index demonstrates factor, construct and longitudinal validity. *BMC Musculoskeletal Disorders, 7*, 12. doi:10.1186/1471-2474-7-12

MacDonald, P. B., Clark, P., & Sutherland, K. (2000). An analysis of the diagnostic accuracy of the Hawkins and Neer subacromial impingement signs. *Journal of Shoulder and Elbow Surgery, 9*(4), 299-301 doi:10.1067/mse.2000.106918

Macfarlane, G. J., Hunt, I. M., & Silman, A. J. (1998). Predictors of chronic shoulder pain: A population based prospective study. *Journal of Rheumatology, 25*(8), 1612-1615.

Maitland, G. D. (1986). *Vertebral Manipulation* (5th ed.). Sydney, Australia: Butterworths.

Mallon, W. J., Herring, C. L., Sallay, P. I., Moorman, C. T., & Crim, J. R. (1996). Use of vertebral levels to measure presumed internal rotation at the shoulder: A radiographic analysis. *Journal of Shoulder and Elbow Surgery, 5*(4), 299-306.

Manchikanti, L., Derby, R., Wolfer, L., Singh, V., Datta, S., & Hirsch, J. A. (2009). Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 5. Diagnostic accuracy studies. *Pain Physician, 12*(3), 517-540.

Manchikanti, L., Dunbar, E. E., Wargo, B. W., Shah, R. V., Derby, R., & Cohen, S. P. (2009). Systematic review of cervical discography as a diagnostic test for chronic spinal pain. *Pain Physician, 12*(2), 305-321.

Manchikanti, L., Glaser, S. E., Wolfer, L., Derby, R., & Cohen, S. P. (2009). Systematic review of lumbar discography as a diagnostic test for chronic low back pain. *Pain Physician, 12*(3), 541-559.

Manchikanti, L., Pampati, S., & Cash, K. A. (2010). Making sense of the accuracy of diagnostic lumbar facet joint nerve blocks: An assessment of the implications of 50% relief, 80% relief, single block, or controlled diagnostic blocks. *Pain Physician, 13*(2), 133-143.

Manchikanti, L., Pampati, V., Fellows, B., & Bakhit, C. E. (1999). Prevalence of lumbar facet joint pain in chronic low back pain. *Pain Physician, 2*(3), 59-64.

Manchikanti, L., Pampati, V., Fellows, B., & Bakhit, C. E. (2000). The diagnostic validity and therapeutic value of lumbar facet joint nerve blocks with or without adjuvant agents. *Current Review of Pain, 4*(5), 337-344.

Martin-Hervas, C., Romero, J., Navas-Acien, A., Reboiras, J. J., & Munuera, L. (2001). Ultrasonographic and magnetic resonance images of rotator cuff lesions compared with arthroscopy or open surgery findings. *Journal of Shoulder and Elbow Surgery, 10*(5), 410-415.

Martinoli, C., Bianchi, S., Prato, N., Pugliese, F., Zamorani, M. P., Valle, M., & Derchi, L. E. (2003). US of the Shoulder: Non-rotator cuff disorders. *Radiographics, 23*(2), 381-401.

Matava, M. J., Purcell, D. B., & Rudzki, J. R. (2005). Partial-thickness rotator cuff tears. *American Journal of Sports Medicine, 33*(9), 1405-1417. doi:10.1177/0363546505280213

Mattingly, G. E., & Mackarey, P. J. (1996). Optimal methods for shoulder tendon palpation: A cadaver study. *Physical Therapy, 76*(2), 166-174.

Mazzocca, A. D., Arciero, R. A., & Bicos, J. (2007). Evaluation and treatment of acromioclavicular joint injuries. *American Journal of Sports Medicine, 35*(2), 316-329. doi:10.1177/0363546506298022

McCabe, R. A., Nicholas, S. J., Montgomery, K. D., Finneran, J. J., McHugh, M. P., McCabe, R. A., ... McHugh, M. P. (2005). The effect of rotator cuff tear size on shoulder strength and range of motion. *Journal of Orthopaedic & Sports Physical Therapy, 35*(3), 130-135.

McFarland, E. G., Hsu, C. Y., Neira, C., & O'Neil, O. (1999). Internal impingement of the shoulder: A clinical and arthroscopic analysis. *Journal of Shoulder and Elbow Surgery, 8*(5), 458-460.

McFarland, E. G., Kim, T. K., & Savino, R. M. (2002). Clinical assessment of three common tests for superior labral anterior-posterior lesions. *American Journal of Sports Medicine, 30*(6), 810-815.

McFarland, E. G., Tanaka, M. J., Garzon-Muvdi, J., Jia, X., & Petersen, S. A. (2009). Clinical and Imaging Assessment for Superior Labrum Anterior and Posterior Lesions. *Current Sports Medicine Reports, 8*(5), 234-239.

McNally, E. G. (2004). *Practical Musculoskeletal Ultrasound* (1st ed.). London: Churchill Livingstone.

Meislin, R. J., Sperling, J. W., & Stitik, T. P. (2005). Persistent shoulder pain: Epidemiology, pathophysiology, and diagnosis. *American Journal of Orthopedics, 34*(12), 5-9.

Meister, K., Buckley, B., & Batts, J. (2004). The posterior impingement sign: Diagnosis of rotator cuff and posterior labral tears secondary to internal impingement in overhand athletes. *American Journal of Orthopedics, 33*(8), 412-415.

Meserve, B. B., Cleland, J. A., & Boucher, T. R. (2009). A meta-analysis examining clinical test utility for assessing superior labral anterior posterior lesions. *The American Journal of Sports Medicine, 37*(11), 2252-2258. doi:10.1177/0363546508325153

Michener, L. A., Walsworth, M. K., Doukas, W. C., & Murphy, K. P. (2009). Reliability and diagnostic accuracy of 5 physical examination tests and combination of tests for subacromial impingement. *Archives of Physical Medicine and Rehabilitation, 90*(11), 1898-1903. doi:10.1016/j.apmr.2009.05.015

Middleton, W. D., Reinus, W. R., & Totty, W. G. (1985). US of the biceps tendon apparatus. *Radiology, 157*(1), 211-215.

Middleton, W. D., Teefey, S. A., & Yamaguchi, K. (2004). Sonography of the rotator cuff: Analysis of interobserver variability. *American Journal of Roentgenology, 183*(5), 1465-1468.

Milgrom, C., Schaffler, M., Gilbert, S., & van Holsbeeck, M. (1995). Rotator-cuff changes in asymptomatic adults. The effect of age, hand dominance and gender. *Journal of Bone and Joint Surgery (British), 77*, 296-298.

Miller, C. A., Forrester, G. A., & Lewis, J. S. (2008a). The validity of the lag signs in diagnosing full-thickness tears of the rotator cuff: a preliminary

investigation [Validation Studies]. *Archives of Physical Medicine & Rehabilitation, 89*(6), 1162-1168.

Miller, C. A., Forrester, G. A., & Lewis, J. S. (2008b). The validity of the lag signs in diagnosing full-thickness tears of the rotator cuff: A preliminary investigation. *Archives of Physical Medicine and Rehabilitation, 89*(6), 1162-1168. doi:10.1016/j.apmr.2007.10.046

Milosavljevic, J., Elvin, A., & Rahme, H. (2005). Ultrasonography of the rotator cuff: A comparison with arthroscopy in one-hundred-and-ninety consecutive cases. *Acta Radiologica, 46*(8), 858-865. doi:10.1080/02841850500270506

Mimori, K., Muneta, T., Nakagawa, T., & Shinomiya, K. (1999). A new pain provocation test for superior labral tears of the shoulder. *American Journal of Sports Medicine, 27*(2), 137-142.

Miniaci, A., Mascia, A. T., Salonen, D. C., & Becker, E. J. (2002). Magnetic resonance imaging of the shoulder in asymptomatic professional baseball pitchers. *American Journal of Sports Medicine, 30*(1), 66-73.

Mintken, P. E., Cleland, J. A., Whitman, J. M., & George, S. Z. (2010). Psychometric properties of the Fear-Avoidance Beliefs Questionnaire and Tampa Scale of Kinesiophobia in patients with shoulder pain. *Archives of Physical Medicine and Rehabilitation, 91*(7), 1128-1136. doi:10.1016/j.apmr.2010.04.009

Miranda, H., Viikari-Juntura, E., Heistaro, S., Heliovaara, M., & Riihimaki, H. (2005). A population study on differences in the determinants of a specific shoulder disorder versus nonspecific shoulder pain without clinical findings. *American Journal of Epidemiology, 161*(9), 847-855. doi:10.1093/aje/kwi112

Mirkovic, M., Green, R., Taylor, N., & Perrott, M. (2005). Accuracy of clinical tests to diagnose superior labral anterior and posterior (SLAP) lesions. *Physical Therapy Reviews, 10*(1), 5-14. doi:10.1179/108331905X43427

Mitchell, C., Adebajo, A., Hay, E., & Carr, A. (2005). Shoulder pain: diagnosis and management in primary care. *British Medical Journal, 331*, 1124-1128. doi:10.1136/bmj.331.7525.1124

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *British Medical Journal, 339*, 332-336. doi:10.1136/bmj.b2535

Monu, J. U. V., Pope Jr, T. L., Chabon, S. J., & Vanarthos, W. J. (1994). MR diagnosis of superior labral anterior posterior (SLAP) injuries of the glenoid labrum: Value of routine imaging without intraarticular injection of contrast material. *American Journal of Roentgenology, 163*(6), 1425-1429.

Moosikasuwan, J. B., Miller, T. T., & Burke, B. J. (2005). Rotator cuff tears: Clinical, radiographic, and US findings. *Radiographics, 25*(6), 1591-1607. doi:10.1148/rg.256045203

Moosmayer, S., Heir, S., & Smith, H. J. (2007). Sonography of the rotator cuff in painful shoulders performed without knowledge of clinical information: Results from 58 sonographic examinations with surgical correlation. *Journal of Clinical Ultrasound, 35*(1), 20-26. doi:10.1002/jcu.20286

Moosmayer, S., & Smith, H. J. (2005). Diagnostic ultrasound of the shoulder - A method for experts only? *Acta Orthopaedica, 76*(4), 503-508. doi:10.1080/17453670510041484

Moosmayer, S., Smith, H. J., Tariq, R., & Larmo, A. (2009). Prevalence and characteristics of asymptomatic tears of the rotator cuff: An ultrasonographic and clinical study. *Journal of Bone and Joint Surgery (British), 91*(2), 196-200. doi:10.1302/0301-620x.91b2.21069

Mulcahy, K. A., Baxter, A. D., Oni, O. O., & Finlay, D. (1994). The value of shoulder distension arthrography with intraarticular injection of steroid and local anaesthetic: A follow-up study. *British Journal of Radiology, 67*(795), 263-266.

Munro, W., & Healy, R. (2009). The validity and accuracy of clinical tests used to detect labral pathology of the shoulder: A systematic review. *Manual Therapy, 14*(2), 119-130. doi:10.1016/j.math.2008.08.008

Murrell, G. A. C., & Walton, J. (2001). Diagnosis of rotator cuff tears. *Lancet, 357*(9258), 769-770.

Myers, T. H., Zemanovic, J. R., & Andrews, J. R. (2005). The resisted supination external rotation test: A new test for the diagnosis of superior labral anterior posterior lesions. *American Journal of Sports Medicine, 33*(9), 1315-1320. doi:10.1177/0363546504273050

Nakagawa, S., Yoneda, M., Hayashida, K., Obata, M., Fukushima, S., & Miyazaki, Y. (2005). Forced shoulder abduction and elbow flexion test: A new simple clinical test to detect superior labral injury in the throwing shoulder. *Arthroscopy, 21*(11), 1290-1295. doi:10.1016/j.arthro.2005.08.025

Nakagawa, Y., Hyakuna, K., Otani, S., Hashitani, M., & Nakamura, T. (1999). Epidemiologic study of glenohumeral osteoarthritis with plain radiography. *Journal of Shoulder and Elbow Surgery, 8*(6), 580-584.

Nanda, R., Gupta, S., Kanapathipillai, P., Liow, R. Y., & Rangan, A. (2008). An assessment of inter examiner reliability of clinical tests for subacromial impingement and rotator cuff integrity. *European Journal of Surgery and Traumatology, 18*, 495-500. doi:10.1007/s00590-008-0341-6

Naredo, E., Aguado, P., De Miguel, E., Uson, J., Mayordomo, L., Gijon-Banos, J., & Martin-Mola, E. (2002). Painful shoulder: Comparison of physical examination and ultrasonographic findings. *Annals of the Rheumatic Diseases, 61*(2), 132-136.

Naredo, E., Aguado, P., Padron, M., Bernad, M., Uson, J., Mayordomo, L., & Martin-Mola, E. (1999). A comparative study of ultrasonography with magnetic resonance imaging in patients with painful shoulder. *Journal of Clinical Rheumatology, 5*(4), 184-192.

Naredo, E., Moller, I., Moragues, C., De Agustin, J. J., Scheel, A. K., Grassi, W., ... Werner, C. (2006). Interobserver reliability in musculoskeletal ultrasonography: Results from a "Teach the Teachers" rheumatologist course. *Annals of the Rheumatic Diseases, 65*(1), 14-19. doi:10.1136/ard.2005.037382

Needell, S. D., & Zlatkin, M. B. (1997). Comparison of fat-saturation fast spin echo versus conventional spin-echo MRI in the detection of rotator cuff pathology. *Journal of Magnetic Resonance Imaging, 7*(4), 674-677.

Neer, C. S. (1972). Anterior acromioplasty for the chronic impingement syndrome in the shoulder: a preliminary report. *Journal of Bone and Joint Surgery (American), 54*(1), 41-50.

Neer, C. S. (1983). Impingement lesions. *Clinical Orthopaedics & Related Research, 173*, 70-78.

Neri, B. R., Chan, K. W., & Kwon, Y. W. (2009). Management of massive and irreparable rotator cuff tears. *Journal of Shoulder and Elbow Surgery, 18*(5), 808-818. doi:10.1016/j.jse.2009.03.013

Norregaard, J., Krogsgaard, M. R., Lorenzen, T., & Jensen, E. M. (2002). Diagnosing patients with longstanding shoulder joint pain. *Annals of Rheumatic Diseases, 61*, 646-649.

O'Brien, S. J., Pagnani, M. J., Fealy, S., McGlynn, S. R., & Wilson, J. B. (1998). The active compression test: A new and effective test for diagnosing labral tears and acromioclavicular joint abnormality. *American Journal of Sports Medicine, 26*(5), 610-613.

O'Connor, P. J., Rankine, J., Gibbon, W. W., Richardson, A., Winter, F., & Miller, J. H. (2005). Interobserver variation in sonography of the painful shoulder. *Journal of Clinical Ultrasound, 33*(2), 53-56. doi:10.1002/jcu.20088

Oh, J. H., Kim, J. Y., Kim, W. S., Gong, H. S., & Lee, J. H. (2008). The evaluation of various physical examinations for the diagnosis of type II superior labrum anterior and posterior lesion. *American Journal of Sports Medicine, 36*(2), 353-359. doi:10.1177/0363546507308363

Olofsson, Y., Book, C., & Jacobsson, L. T. H. (2003). Shoulder joint involvement in patients with newly diagnosed rheumatoid arthritis. *Scandinavian Journal of Rheumatology, 32*(1), 25-32. doi:10.1080/03009740310000364

Ombregt, L., Bisschop, P., & ter Veer, H. J. (2003). Clinical examination of the shoulder. In L. Ombregt (Ed.), *System of Orthopaedic Medicine* (pp. 291-302). London: Churchill Livingstone.

Ostor, A. J. K., Richards, C. A., Prevost, A. T., Hazleman, B. L., & Speed, C. A. (2004). Inter-rater reliability of clinical tests for rotator cuff lesions. *Annals of Rheumatic Disorders, 63*, 1288-1292. doi:10.1136/ard.2003.014712

Ostor, A. J. K., Richards, C. A., Prevost, A. T., Speed, C. A., & Hazleman, B. L. (2005). Diagnosis and relation to general health of shoulder disorders

presenting to primary care. *Rheumatology, 44*, 800-805. doi:10.1093/rheumatology/keh598

Ottenheijm, R. P., Jansen, M. J., Staal, J. B., Van Den Bruel, A., Weijers, R. E., De Bie, R. A., & Dinant, G. J. (2010). Accuracy of diagnostic ultrasound in patients with suspected subacromial disorders: A systematic review and meta-analysis. *Archives of Physical Medicine and Rehabilitation, 91*(10), 1616-1625. doi:10.1016/j.apmr.2010.07.017

Paavolainen, P., & Ahovuo, J. (1994). Ultrasonography and arthrography in the diagnosis of tears of the rotator cuff. *Journal of Bone and Joint Surgery (American), 76*(3), 335-340.

Palmer, K., Walker-Bone, K., Linaker, C., Reading, I., Kellingray, S., Coggon, D., & Cooper, C. (2000). The Southampton examination schedule for the diagnosis of musculoskeletal disorders of the upper limb. *Annals of Rheumatic Disorders, 59*, 5-11.

Palmer, W. E., & Caslowitz, P. L. (1995). Anterior shoulder instability: Diagnostic criteria determined from prospective analysis of 121 MR arthrograms. *Radiology, 197*(3), 819-825.

Pandya, N. K., Colton, A., Webner, D., Sennett, B., & Huffman, G. R. (2008). Physical examination and magnetic resonance imaging in the diagnosis of superior labrum anterior-posterior lesions of the shoulder: A sensitivity analysis. *Arthroscopy, 24*(3), 311-317. doi:10.1016/j.arthro.2007.09.004

Parentis, M. A., Glousman, R. E., Mohr, K. S., & Yocum, L. A. (2006). An evaluation of the provocative tests for superior labral anterior posterior lesions. *American Journal of Sports Medicine, 34*(2), 265-268. doi:10.1177/0363546505279911

Park, H. B., Yokota, A., Gill, H. S., el Rassi, G., & McFarland, E. G. (2005). Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndrome. *Journal of Bone and Joint Surgery (American), 87*(7), 1446-1455.

Partington, P. F., & Broome, G. H. (1998). Diagnostic injection around the shoulder: Hit and miss? A cadaveric study of injection accuracy. *Journal of Shoulder and Elbow Surgery, 7*(2), 147-150.

Patte, D. (1990). Classification of rotator cuff lesions. *Clinical Orthopaedics and Related Research, 254*, 81-86.

Peat, J., & Barton, B. (2005). Categorical variables: risk statistics. In M. Banks (Ed.), *Medical Statistics: A guide to data analysis and critical appraisal* (pp. 241-266). Massachusetts: Blackwell Publishing Inc.

Petersson, C. J. (1986). Painful shoulders in patients with rheumatoid arthritis. Prevalence, clinical and radiological features. *Scandinavian Journal of Rheumatology, 15*(3), 275-279.

Pfirrmann, C. W. A., Zanetti, M., Weishaupt, D., Gerber, C., & Hodler, J. (1999). Subscapularis tendon tears: Detection and grading at MR arthrography. *Radiology, 213*(3), 709-714.

Plafki, C., Steffen, R., E., W. R., & Wittenberg, R. H. (2000). Local anaesthetic injection with and without corticosteroids for subacromial impingement syndrome. *International Orthopaedics, 24*(1), 40-42.

Pope, D. P., Croft, P. R., Pritchard, C. M., & Silman, A. J. (1997). Prevalence of shoulder pain in the community: The influence of case definition. *Annals of the Rheumatic Diseases, 56*(5), 308-312.

Powell, J. W., Huijbregts, P. A., & Jensen, R. (2008). Diagnostic utility of clinical tests for SLAP lesions: A systematic literature review. *Journal of Manual and Manipulative Therapy, 16*(3), E58-E59.

Pradhan, R. L., & Itoi, E. (2001). Rotator interval lesions of the shoulder joint. *Orthopedics, 24*(8), 798-801.

Precerutti, M., Garioni, E., Madonia, L., & Draghi, F. (2010). US anatomy of the shoulder: Pictorial essay. *Journal of Ultrasound, 13*(4), 179-187. doi:10.1016/j.jus.2010.10.005

Pyne, S. W. (2004). Diagnosis and current treatment options of shoulder impingement. *Current Sports Medicine Reports., 3*(5), 251-255.

R Development Core Team. (2010). R: A language and environment for statistical computing [computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org.

Rawlings, R. R., Rae, D. S., & Graubard, B. I. (1982). A methodology for construction of a multivariate diagnostic instrument: An application to alcohol abuse screening. *Computers and Biomedical Research, 15*(3), 228-239.

Razmjou, H., Holtby, R., & Myhr, T. (2004). Pain provocative shoulder tests: Reliability and validity of the impingement tests. *Physiotherapy Canada, 56*(4), 229-236.

Read, J. W., & Perko, M. (1998). Shoulder ultrasound: Diagnostic accuracy for impingement syndrome, rotator cuff tear, and biceps tendon pathology. *Journal of Shoulder and Elbow Surgery, 7*, 264-271.

Reid, M. C., Lachs, M. S., & Feinstein, A. R. (1995). Use of methodological standards in diagnostic test research: Getting better but still not good. *Journal of the American Medical Association, 274*(8), 645-651.

Reilly, P., Macleod, I., Macfarlane, R., Windley, J., & Emery, R. J. H. (2006). Dead men and radiologists don't lie: A review of cadaveric and radiological studies of rotator cuff tear prevalence. *Annals of the Royal College of Surgeons of England, 88*(2), 116-121. doi:10.1308/003588406X94968

Revel, M., Poiraudeau, S., Auleley, G. R., Payan, C., Denke, A., Nguyen, M., ... Fermanian, J. (1998). Capacity of the clinical picture to characterize low

back pain relieved by facet joint anesthesia. Proposed criteria to identify patients with painful facet joints. *Spine, 23*(18), 1972-1977. doi:10.1097/00007632-199809150-00011

Riddle, D. L., Rothstein, J. M., & Lamb, R. L. (1987). Goniometric reliability in a clinical setting. Shoulder measurements. *Physical Therapy, 67*(5), 668-673.

Rigby, A. S. (2000). Statistical methods in epidemiology. V. Towards an understanding of the kappa coefficient. *Disability & Rehabilitation, 22*(8), 339-344.

Roach, K. E., Budiman-Mak, E., Songsiridej, N., & Lertratanakul, Y. (1991). Development of a shoulder pain and disability index. *Arthritis Care and Research, 4*(4), 143-149.

Robertson, P. L., Schweitzer, M. E., Mitchell, D. G., Schlesinger, F., Epstein, R. E., Frieman, B. G., & Fenlin, J. M. (1995). Rotator cuff disorders: Interobserver and intraobserver variation in diagnosis with MR imaging. *Radiology, 194*(3), 831-835.

Rockwood, C. A., Williams, G., & Young, D. (1998). Disorders of the acromioclavicular joint. In C. A. Rockwood, F. A. Matsen III, M. A. Wirth, & D. T. Harryman II (Eds.), *The Shoulder* (pp. 483-553). Philadelphia: WB Saunders.

Rutjes, A. W. S., Reitsma, J. B., Di Nisio, M., Smidt, N., Van Rijn, J. C., & Bossuyt, P. M. M. (2006). Evidence of bias and variation in diagnostic accuracy studies. *Canadian Medical Association Journal, 174*(4), 469-476. doi:10.1503/cmaj.050090

Rutten, M. J. C., Jager, G. J., & Blickman, J. G. (2006). US of the rotator cuff: Pitfalls, limitations, and artifacts. *RadioGraphics, 26*, 589-604. doi:10.1148/rg.262045719

Rutten, M. J. C., Maresch, B. J., Jager, G. J., & Malefijt, M. C. D. W. (2007). Injection of the subacromial-subdeltoid bursa: Blind or ultrasound-guided? *Acta Orthopaedica Scandinavica, 78*(2), 254-257. doi:10.1080/17453670710013762

Sackett, D. L. (1983). Interpretation of diagnostic data: 5. How to do it with simple maths. *Canadian Medical Association Journal, 129*(9), 947-954.

Sackett, D. L. (1992). A primer on the precision and accuracy of the clinical examination. *Journal of the American Medical Association, 267*(19), 2638-2644.

Şahin, G., & Demirtaş, M. (2006). An overview of MR arthrography with emphasis on the current technique and applicational hints and tips. *European Journal of Radiology, 58*(3), 416-430. doi:10.1016/j.ejrad.2006.01.002

Salzman, K. L., Lillegard, W. A., & Butcher, J. D. (1997). Upper extremity bursitis. *American Family Physician, 56*(7), 1797-1806.

Sammarco, V. J. (2000). Os acromiale: Frequency, anatomy, and clinical implications. *Journal of Bone & Joint Surgery, 82*(3), 394-400.

Santavirta, S., Konttinen, Y. T., Antti-Poika, I., & Nordstrom, D. (1992). Inflammation of the subacromial bursa in chronic shoulder pain. *Archives of Orthopaedic and Trauma Surgery, 111*(6), 336-340. doi:10.1007/bf00420062

Sarkar, K., & Uhthoff, H. K. (1983). Ultrastructure of the subacromial bursa in painful shoulder syndromes. *Virchows Archiv, 400*(2), 107-117.

Schaardenburg, D., van den Brande, K. J. S., Ligthart, G. J., Breedveld, F. C., & Hazes, J. M. W. (1994). Musculokeletal disorders in persons aged 85 and over, a communitiy survey. *Annals of Rheumatic Diseases, 53*, 807-811.

Scheel, A. K., Schmidt, W. A., Hermann, K. G. A., Bruyn, G. A., D'Agostino, M. A., Grassi, W., ... Backhaus, M. (2005). Interobserver reliability of rheumatologists performing musculoskeletal ultrasonography: Results from a EULAR "Train the trainers" course. *Annals of the Rheumatic Diseases, 64*(7), 1043-1049. doi:10.1136/ard.2004.030387

Scheibel, M., Magosch, P., Pritsch, M., Lichtenberg, S., & Habermeyer, P. (2005). The Belly-Off sign: A new clinical diagnostic sign for subscapularis lesions. *Arthroscopy, 21*(10), 1229-1235. doi:10.1016/j.arthro.2005.06.021

Schellingerhout, J. M., Verhagen, A. P., Thomas, S., & Koes, B. W. (2008). Lack of uniformity in diagnostic labeling of shoulder pain: Time for a different approach. *Manual Therapy, 13*, 478-483. doi:10.1016/j.math.2008.04.005

Schlechter, J. A., Summa, S., & Rubin, B. D. (2009). The passive distraction test: A new diagnostic aid for clinically significant superior labral pathology. *Arthroscopy, 25*(12), 1374-1379. doi:10.1016/j.arthro.2009.04.070

Schmidt, W. A., Schmidt, H., Schicke, B., & Gromnica-Ihle, E. (2004). Standard reference values for musculoskeletal ultrasonography. *Annals of the Rheumatic Diseases, 63*(8), 988-994. doi:10.1136/ard.2003.015081

Schreinemachers, S. A., Van Der Hulst, V. P. M., Willems, W. J., Bipat, S., & Van Der Woude, H. J. (2009). Is a single direct MR arthrography series in ABER position as accurate in detecting anteroinferior labroligamentous lesions as conventional MR arthography? *Skeletal Radiology, 38*(7), 675-683. doi:10.1007/s00256-009-0692-z

Schwartz, A. (2002). *Diagnostic test calculator (version 2006032401)*. Retrieved from http://araw.mede.uic.edu/cgialansz/testcalc.pl?DT=&Dt=&dT=&dt=&2x2=Compute

Schwarzer, A. C., April, C. N., Derby, R., Fortin, J., Kine, G., & Bogduk, N. (1994). The false-positive rate of uncontrolled diagnostic blocks of the lumbar zygapophysial joints. *Pain, 58*(2), 195-200. doi:10.1016/0304-3959(94)90199-6

Schwarzer, A. C., April, C. N., Derby, R., Fortin, J., Kine, G., & Bogduk, N. (1995). The prevalence and clinical features of internal disc disruption in patients with chronic low back pain. *Spine, 20*(17), 1878-1883.

Schwarzer, A. C., Derby, R., April, C. N., Fortin, J., Kine, G., & Bogduk, N. (1994). The value of the provocation response in lumbar zygapophyseal joint injections. *Clinical Journal of Pain, 10*(4), 309-313.

Shaffer, B. S. (1999). Painful conditions of the acromioclavicular joint. *The Journal of the American Academy of Orthopaedic Surgeons, 7*(3), 176-188.

Shahabpour, M., Kichouh, M., Laridon, E., Gielen, J. L., & De Mey, J. (2008). The effectiveness of diagnostic imaging methods for the assessment of soft tissue and articular disorders of the shoulder and elbow. *European Journal of Radiology, 65*(2), 194-200. doi:10.1016/j.ejrad.2007.11.012

Shankar, V., & Bangdiwala, S. I. (2008). Behaviour of agreement measures in the presence of zero cells and biased marginal distributions. *Journal of Applied Statistics, 35*(4), 445-464. doi:10.1080/02664760701835052

Sheps, S. B., & Schechter, M. T. (1984). The assessment of diagnostic tests. A survey of current medical research. *Journal of the American Medical Association, 252*(17), 2418-2422. doi:10.1001/jama.252.17.2418

Sher, J. S., Iannotti, J. P., Williams, G. R., Herzog, R. J., Kneeland, J. B., Lisser, S., & Patel, N. (1998). The effect of shoulder magnetic resonance imaging on clinical decision making. *Journal of Shoulder and Elbow Surgery, 7*(3), 205-209.

Sher, J. S., Uribe, J. W., Posada, A., Murphy, B. J., & Zlatkin, M. B. (1995). Abnormal findings on magnetic resonance images of asymptomatic shoulders. *Journal of Bone and Joint Surgery, 77*(1), 10-15.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.

Shubin Stein, B. E., Ahmad, C. S., Pfaff, C. H., Bigliani, L. U., & Levine, W. N. (2006). A comparison of magnetic resonance imaging findings of the acromioclavicular joint in symptomatic versus asymptomatic patients. *Journal of Shoulder and Elbow Surgery, 15*(1), 56-59. doi:10.1016/j.jse.2005.05.013

Shubin Stein, B. E., Wiater, J. M., Pfaff, H. C., Bigliani, L. U., & Levine, W. N. (2001). Detection of acromioclavicular joint pathology in asymptomatic shoulders with magnetic resonance imaging. *Journal of Shoulder and Elbow Surgery, 10*(3), 204-208. doi:10.1067/mse.2001.113498

Skedros, J. G., Hunt, K. J., & Pitts, T. C. (2007). Variations in corticosteroid/anesthetic injections for painful shoulder conditions: comparisons among orthopaedic surgeons, rheumatologists, and physical medicine and primary-care physicians. *BMC Musculoskeletal Disorders, 8*, 63. doi:10.1186/1471-2474-8-63

Smidt, N., & Green, S. (2003). Is the diagnosis important for the treatment of patients with shoulder complaints? *Lancet, 362*(9399), 1867-1868.

Snyder, S. J. (1993). Evaluation and treatment of the rotator cuff. *Orthopedic Clinics of North America, 24*(1), 173-192.

Snyder, S. J., Banas, M. P., & Karzel, R. P. (1995). An analysis of 140 injuries to the superior glenoid labrum. *Journal of Shoulder and Elbow Surgery, 4*(4), 243-248.

Snyder, S. J., Pachelli, A. F., Del Pizzo, W., Friedman, M. J., Ferkel, R. D., & Pattee, G. (1991). Partial thickness rotator cuff tears: Results of arthroscopic treatment. *Arthroscopy: The Journal of Arthroscopic & Related Surgery, 7*(1), 1-7. doi:10.1016/0749-8063(91)90070-e

Solomon, D. H., Bates, D. W., Schaffer, J. L., Horsky, J., Burdick, E., & Katz, J. N. (2001). Referrals for musculoskeletal disorders: patterns, predictors, and outcomes. *Journal of Rheumatology, 28*(9), 2090-2095.

Sonnabend, D. H., Hughes, J. S., Giuffre, B. M., & Farrell, R. (1997). The clinical role of shoulder ultrasound. *Australian and New Zealand Journal of Surgery, 67*(9), 630-633.

SooHoo, N. F., & Rosen, P. (1996). Diagnosis and treatment of rotator cuff tears in the emergency department. *Journal of Emergency Medicine, 14*(3), 309-317. doi:10.1016/0736-4679(96)00027-3

Speed, C. A., & Hazleman, B. L. (1999). Calcific tendinitis of the shoulder. *New England Journal of Medicine, 340*(20), 1582-1584.

Steinbach, L. S., Palmer, W. E., & Schweitzer, M. E. (2002). Special focus session: MR arthrography. *Radiographics, 22*(5), 1223-1246.

Stetson, W. B., & Templin, K. (2002). The Crank test, the O'Brien test, and routine magnetic resonance imaging scans in the diagnosis of labral tears. *American Journal of Sports Medicine, 30*(6), 806-809.

Stoller, D. W., Wolf, E. M., Li, A. E., Nottage, W. M., & Tirman, P. F. J. (2007). The Shoulder. In D. W. Stoller (Ed.), *Magnetic Resonance Imaging in Orthopaedics and Sports Medicine* (pp. 1131-1462). Baltimore: Lippincott Williams and Wilkins.

Strobel, K., Pfirrmann, C. W. A., Zanetti, M., Nagy, L., & Hodler, J. (2003). MRI features of the acromioclavicular joint that predict pain relief from intraarticular injection. *American Journal of Roentgenology, 181*, 755-760.

Taljanovic, M. S., Carlson, K. L., Kuhn, J. E., Jacobson, J. A., Delaney-Sathy, L. O., & Adler, R. S. (2000). Sonography of the glenoid labrum: A cadaveric study with arthroscopic correlation. *American Journal of Roentgenology, 174*(6), 1717-1722.

Teefey, S. A., Rubin, D. A., Middleton, W. D., Hildebolt, C. F., Leibold, R. A., & Yamaguchi, K. (2004). Detection and quantification of rotator cuff tears: Comparison of ultrasonographic, magnetic resonance imaging and

arthroscopic findings in seventy one consecutive cases. *Journal of Bone and Joint Surgery, 86*(4), 708-716.

Tempelhof, S., Rupp, S., & Seil, R. (1999). Age-related prevalence of rotator cuff tears in asymptomatic shoulders. *Journal of Shoulder and Elbow Surgery, 8*(4), 296-299.

Tennent, T. D., Beach, W. R., & Meyers, J. F. (2003). A review of the special tests associated with shoulder examination. Part 1: The rotator cuff tests. *American Journal of Sports Medicine, 31*(1), 154-160.

Terwee, C. B., de Winter, A. F., Scholten, J. P., Jans, M. P., Deville, W., van Schaardenburg, D., & Bouter, L. M. (2005). Interobserver reproducibility of the visual estimation of range of motion of the shoulder. *Archives of Physical Medicine and Rehabilitation, 86*(7), 1356-1361. doi:10.1016/j.apmr.2004.12.031

Thornbury, J. R. (1994). Clinical efficacy of diagnostic imaging: Love it or leave it. *American Journal of Roentgenology, 162*(1), 1-8.

Tokish, J. M., Decker, M. J., Ellis, H. B., Torry, M. R., & Hawkins, R. J. (2003). The belly-press test for the physical examination of the subscapularis muscle: Electromyographic validation and comparison to the lift-off test. *Journal of Shoulder and Elbow Surgery, 12*(5), 427-430. doi:10.1016/S1058-2746(03)00047-8

Traughber, P. D., & Goodwin, T. E. (1992). Shoulder MRI: Arthroscopic correlation with emphasis on partial tears. *Journal of Computer Assisted Tomography, 16*(1), 129-133.

Triffitt, P. D., Wildin, C., & Hajioff, D. (1999). The reproducibility of measurement of shoulder movement. *Acta Orthopedica Scandinavica, 70*, 322-324.

Tsai, Y. H., Huang, T. J., Hsu, W. H., Huang, K. C., Li, Y. Y., Peng, K. T., & Hsu, R. W. W. (2007). Detection of subacromial bursa thickening by sonography in shoulder impingement syndrome. *Chang Gung Medical Journal, 30*(2), 135-141.

Tuite, M. J., Rutkowski, A., Enright, T., Kaplan, L., Fine, J. P., & Orwin, J. (2005). Width of high signal and extension posterior to biceps tendon as signs of superior labrum anterior to posterior tears on MRI and MR arthrography. *American Journal of Roentgenology, 185*, 1422-1428.

Turner-Bowker, D. M., Bayliss, M. S., & Ware, J. E. J. (2003). Usefulness of the SF-8 health survey for comparing the impact of migraine and other conditions. *Quality of Life Research, 12*, 1003-1012.

Turner-Stokes, L. (1996). Clinical differential diagnosis of shoulder pain. *British Journal of Hospital Medicine, 56*(2-3), 73-77.

Tyler, T. F., Nahow, R. C., Nicholas, S. J., McHugh, M. P., Tyler, T. F., Nahow, R. C., ... McHugh, M. P. (2005). Quantifying shoulder rotation weakness in patients with shoulder impingement. *Journal of Shoulder & Elbow Surgery, 14*(6), 570-574. doi:10.1016/j.jse.2005.03.003

Urwin, M., Symmons, D., Allison, T., Brammah, T., Busby, H., Roxby, M., ... Williams, G. (1998). Estimating the burden of musculoskeletal disorders in the community: The comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. *Annals of the Rheumatic Diseases, 57*, 649-655.

van der Windt, D. A. W. M., Koes, B. W., Bocke, A. J. P., Deville, W., de Jong, B. A., & Bounter, L. M. (1996). Shoulder disorders in general practice: Prognostic indicators of outcome. *British Journal of General Practice, 46*, 519-523.

van der Windt, D. A. W. M., Koes, B. W., De Jong, B. A., & Bouter, L. M. (1995). Shoulder disorders in general practice: Incidence, patient characteristics, and management. *Annals of the Rheumatic Diseases, 54*(12), 959-964.

van der Windt, D. A. W. M., Koes, B. W., Devillé, W., Boeke, A. J. P., de Jong, B. A., & Bouter, L. M. (1998). Effectiveness of corticosteroid injections versus physiotherapy for treatment of painful stiff shoulder in primary care: randomised trial. *British Medical Journal, 317*, 1292-1296.

Van Holsbeeck, M., & Strouse, P. J. (1993). Sonography of the shoulder: Evaluation of the subacromial-subdeltoid bursa. *American Journal of Roentgenology, 160*(3), 561-564.

Van Holsbeeck, M. T., Kolowich, P. A., Eyler, W. R., Craig, J. G., Shirazi, K. K., Habra, G. K., ... Bouffard, J. A. (1995). US depiction of partial-thickness tear of the rotator cuff. *Radiology, 197*(2), 443-446.

Van Riet, R. P., & Bell, S. N. (2011). Clinical evaluation of acromioclavicular joint pathology: Sensitivity of a new test. *Journal of Shoulder and Elbow Surgery, 20*(1), 73-76. doi:10.1016/j.jse.2010.05.023

Vangsness, C. T., Ennis, M., Taylor, J. G., & Atkinson, R. (1995). Neural anatomy of the glenohumeral ligaments, labrum, and subacromial bursa. *Arthroscopy, 11*(2), 180-184.

Vecchio, P., Kavanagh, R., Hazleman, B. L., & King, R. H. (1995). Shoulder pain in a community-based rheumatology clinic. *British Journal of Rheumatology, 34*(5), 440-442.

Voloshin, I., Gelinas, J., Maloney, M. D., O'Keefe, R. J., Bigliani, L. U., & Blaine, T. A. (2005). Proinflammatory cytokines and metalloproteases are expressed in the subacromial bursa in patients with rotator cuff disease. *Arthroscopy, 21*(9), 1076.e1071-1076.e1079. doi:10.1016/j.arthro.2005.05.017

Waddell, G., Bircher, M., Finlayson, D., & Main, C. J. (1984). Symptoms and signs: Physical disease or illness behaviour? *British Medical Journal, 289*(6447), 739-741.

Waddell, G., Newton, M., Henderson, I., Somerville, D., & Main, C. J. (1993). A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain, 52*(2), 157-168.

Wadsworth, C. T., Nielsen, D. H., Corcoran, D. S., Phillips, C. E., & Sannes, T. L. (1992). Interrater reliability of hand-held dynamometry: Effects of rater gender, weight, and grip strength. *Journal of Orthopaedic & Sports Physical Therapy, 16*(2), 74-81.

Wagner, S. C., Schweitzer, M. E., Morrison, W. B., Fenlin J.M, Jr., & Bartolozzi, A. R. (2002). Shoulder instability: Accuracy of MR imaging performed after surgery in depicting recurrent injury. *Radiology, 222*(1), 196-203.

Walch, G., Boileau, P., Noel, E., & Donell, S. T. (1992). Impingement of the deep surface of the supraspinatus tendon on the posterosuperior glenoid rim: An arthroscopic study. *Journal of Shoulder and Elbow Surgery, 1*(5), 238-245.

Walch, G., Boulahia, A., Calderone, S., & Robinson, A. (1998). The 'dropping' and 'hornblowers' signs in evaluation of rotator cuff tears. *Journal of Bone and Joint Surgery (British), 80*(4), 624-628.

Waldt, S., Bruegel, M., Mueller, D., Holzapfel, K., Imhoff, A. B., Rummeny, E. J., & Woertler, K. (2007). Rotator cuff tears: Assessment with MR arthrography in 275 patients with arthroscopic correlation. *European Radiology, 17*(2), 491-498. doi:10.1007/s00330-006-0370-7

Waldt, S., Burkart, A., Imhoff, A. B., Bruegel, M., Rummeny, E. J., & Woertler, K. (2005). Anterior shoulder instability: Accuracy of MR arthrography in the classification of anteroinferior labroligamentous injuries. *Radiology, 237*(2), 578-583. doi:10.1148/radiol.2372041429

Waldt, S., Burkart, A., Lange, P., Imhoff, A. B., Rummeny, E. J., & Woertler, K. (2004). Diagnostic performance of MR arthrography in the assessment of superior labral anteroposterior lesions of the shoulder. *American Journal of Roentgenology, 182*, 1271-1278.

Walker, B. F. (2000). The prevalence of low back pain: A systematic review of the literature from 1966 to 1998. *Journal of Spinal Disorders, 13*(3), 205-217. doi:10.1097/00002517-200006000-00003

Walmsley, S., Rivett, D. A., & Osmotherly, P. G. (2009). Adhesive capsulitis: Establishing consensus on clinical identifiers for stage 1 using the DELPHI technique. *Physical Therapy, 89*(9), 906-917.

Walsworth, M. K., Doukas, W. C., Murphy, K. P., Mielcarek, B. J., & Michener, L. A. (2008). Reliability and diagnostic accuracy of history and physical examination for diagnosing glenoid labral tears. *American Journal of Sports Medicine, 36*(1), 162-168. doi:10.1177/0363546507307508

Walton, J., Mahajan, S., Paxinos, A., Marshall, J., Bryant, C., Shnier, R., ... Murrell, G. A. C. (2004). Diagnostic values of tests for acromioclavicular joint pain. *Journal of Bone and Joint Surgery (American), 86*(4), 807-812.

Ware, J. E. J., Kosinski, M., Dewey, J. E., & Gandek, B. (2001). *How to Score and Interpret Single-Item Health Status Measures: A Manual for Users of the SF-8™ Health Survey* (3rd ed.). Lincoln, RI: QualityMetric Incorporated.

Warner, J. J. P., Micheli, L. J., Arslanian, L. E., Kennedy, J., & Kennedy, R. (1990). Patterns of flexibility, laxity, and strength in normal shoulders and shoulders with instability and impingement. *American Journal of Sports Medicine, 18*(4), 366-375.

Wasserlauf, B. L., & Matava, M. J. (2003). Advances in surgical management of superior labrum, anterior-to-posterior lesions. *Current Opinion in Orthopaedics, 14*(4), 262-268. doi:10.1097/00001433-200308000-00009

Wasson, J. H., Sox, H. C., Neff, R. K., & Goldman, L. (1985). Clinical prediction rules. Applications and methodological standards. *New England Journal of Medicine, 313*(13), 793-799.

Werner, C. M. L., Blumenthal, S., Curt, A., & Gerber, C. (2006). Subacromial pressures in vivo and effects of selective experimental suprascapular nerve block. *Journal of Shoulder and Elbow Surgery, 15*(3), 319-323. doi:10.1016/j.jse.2005.08.017

Whiting, P., Harbord, R., & Kleijnen, J. (2005). No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology, 5*. doi:10.1186/1471-2288-5-19

Whiting, P., Rutjes, A. W. S., Dinnes, J., Reitsma, J. B., Bossuyt, P. M. M., & Kleijnen, J. (2005). A systematic review finds that diagnostic reviews fail to incorporate quality despite available tools. *Journal of Clinical Epidemiology, 58*(1), 1-12. doi:10.1016/j.jclinepi.2004.04.008

Whiting, P., Rutjes, A. W. S., Reitsma, J. B., Bossuyt, P. M. M., & Kleijnen, J. (2003). The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology, 3*, 1-13. doi:10.1186/1471-2288-3-1

Whiting, P., Rutjes, A. W. S., Reitsma, J. B., Glas, A. S., Bossuyt, P. M. M., & Kleijnen, J. (2004). Sources of variation and bias in studies of diagnostic accuracy: A systematic review. *Annals of Internal Medicine, 140*(3), 189-202.

WHO Scientific Group. (2000). *The Burden of Musculoskeletal Conditions at the Start of the New Millennium*. Geneva: World Health Organisation.

Wilk, K. E., Reinold, M. M., Dugas, J. R., Arrigo, C. A., Moser, M. W., & Andrews, J. R. (2005). Current concepts in the recognition and treatment of superior labral (SLAP) lesions. *Journal of Orthopaedic and Sports Physical Therapy, 35*(5), 273-291.

Winters, J. C., Sobel, J. S., Groenier, K. H., Arendzen, H. J., & Meyboom-De Jong, B. (1997). Comparison of physiotherapy, manipulation, and corticosteroid injection for treating shoulder complaints in general practice: Randomised, single blind study. *British Medical Journal, 314*(7090), 1320-1325.

Winters, J. C., Sobel, J. S., Groenier, K. H., Arendzen, J. H., & Meyboom - de Jong, B. (1999). The long-term course of shoulder complaints: A prospective study in general practice. *Rheumatology, 38*, 160-163.

Wnorowski, D. C., Levinsohn, E. M., Chamberlain, B. C., & McAndrew, D. L. (1997). Magnetic resonance imaging assessment of the rotator cuff: Is it really accurate? *Arthroscopy, 13*(6), 710-719. doi:10.1016/s0749-8063(97)90005-0

Wright, A. P., MacLeod, I. A. R., & Talwalker, S. C. (2011). Disorders of the acromioclavicular joint and distal clavicle. *Orthopaedics and Trauma, 25*(1), 30-36. doi:10.1016/j.mporth.2010.10.011

Wright, V., & Haq, A. M. M. M. (1976). Periarthritis of the shoulder. I. Aetiological considerations with particular reference to personality factors. *Annals of the Rheumatic Diseases, 35*(3), 213-219.

Yağcı, B., Manisalı, M., Yilmaz, E., Özkan, M., Ekin, A., Özaksoy, D., & Kovanlıkaya, I. (2001). Indirect MR arthrography of the shoulder in detection of rotator cuff ruptures. *European Radiology, 11*(2), 258-262.

Yamaguchi, K., Tetro, A. M., Blam, O., Evanoff, B. A., Teefey, S. A., & Middleton, W. D. (2001). Natural history of asymptomatic rotator cuff tears: A longitudinal analysis of asymptomatic tears detected sonographically. *Journal of Shoulder and Elbow Surgery, 10*(3), 199-203.

Yamakado, K. (2002). The targeting accuracy of subacromial injection to the shoulder: An arthrographic evaluation. *Arthroscopy, 18*(8), 887-891.

Yamanaka, K., & Matsumoto, T. (1994). The joint side tear of the rotator cuff. A followup study by arthrography. *Clinical Orthopaedics and Related Research, 304*, 68-73.

Yen, C. H., Chiou, H. J., Chou, Y. H., Hsu, C. C., Wu, J. J., Ma, H. L., & Chang, C. Y. (2004). Six surgery-correlated sonographic signs for rotator cuff tears: Emphasis on partial-thickness tear. *Clinical Imaging, 28*(1), 69-76. doi:10.1016/s0899-7071(03)00034-2

Zaslav, K. R. (2001). Internal rotation resistance strength test: A new diagnostic test to differentiate intra-articular pathology from outlet (Neer) impingement syndrome in the shoulder. *Journal of Shoulder and Elbow Surgery, 10*(1), 23-27. doi:10.1067/mse.2001.111960

Zehetgruber, H., Lang, T., & Wurnig, C. (2002). Distinction between supraspinatus, infraspinatus and subscapularis tendon tears with ultrasound in 332 surgically confirmed cases. *Ultrasound in Medicine and Biology, 28*(6), 711-717. doi:10.1016/s0301-5629(02)00508-2

Ziegler, D. W. (2004). The use of in-office, orthopaedist-performed ultrasound of the shoulder to evaluate and manage rotator cuff disorders. *Journal of Shoulder and Elbow Surgery, 13*(3), 291-297. doi:10.1016/j.jse.2004.01.017

Zingg, P. O., Jost, B., Sukthankar, A., Buhler, M., Pfirrmann, C. W. A., & Gerber, C. (2007). Clinical and structural outcomes of nonoperative management of massive rotator cuff tears. *Journal of Bone and Joint Surgery (American), 89*(9), 1928-1934. doi:10.2106/jbjs.f.01073

Zuckerman, J. D., Cuomo, F., & Rokito, S. (1994). Definition and classification of frozen shoulder. *Journal of Shoulder and Elbow Surgery, 3*, S72.

APPENDICES

Appendix 1

Ethical Approval

**Health
and
Disability
Ethics
Committees**

Upper South A Regional Ethics Committee
Ministry of Health
4th Floor, 250 Oxford Tce
PO Box 3877
Christchurch
Phone (03) 372 3037
Fax (03) 372 1015
Email: uppersouth_ethicscommittee@moh.govt.nz

9 May 2008

Angela Cadogan
2/34 Devon Street
Beckenham
Christchurch 8023

Dear Angela Cadogan,

Diagnostic accuracy of a clinical examination in determining the source of shoulder pain at primary care level
Investigators: A Cadogan, Dr M Laslett, A/Prof W Hing (Supervisor), A/Prof P McNair, Dr M Coates, Dr T Page
Locality: Primary Care General Medical Practice, Physiotherapy private practices
Ethics ref: URA/08/02/013

The above study has been given ethical approval by the **Upper South A Regional** Ethics Committee. A list of members of this committee is attached.

The Chairperson wishes to pass on her thanks for your attention to detail when compiling your response to the committee's comments.

Approved Documents
Referrer Information sheet dated 30 April 2008
Primary Care Referrer Agreement
Participant record sheet
Reception information
Participant information sheet dated 30 April 2008
Participant consent forms (1) and (2) dated 30 April 2008
Participant information sheet – Reliability Study dated 30 April 2008
Participant consent form – Reliability Study dated 30 April 2008

Please note that the approved documents incorporate the amended methodology as described in the document 'Amendment to Methodology' dated 30 April 2008.

Certification
The Committee is satisfied that this study is not being conducted principally for the benefit of the manufacturer or distributor of the medicine or item in respect of which the trial is being carried out.

Accreditation
The Committee involved in the approval of this study is accredited by the Health Research Council and is constituted and operates in accordance with the Operational Standard for Ethics Committees, April 2006.

Progress Reports
The study is approved until 28 February 2012. The Committee will review the approved application annually and notify the Principal Investigator if it withdraws approval. It is the Principal Investigator's responsibility to forward a progress report covering all sites prior to ethical review of the project in May

2009. The report form is available on http://www.ethicscommittees.health.govt.nz. Please note that failure to provide a progress report may result in the withdrawal of ethical approval. A final report is also required at the conclusion of the study.

Requirements for SAE Reporting

The Principal Investigator will inform the Committee as soon as possible of the following:

- Any related study in another country that has stopped due to serious or unexpected adverse events
- withdrawal from the market for any reason
- all serious adverse events occurring during the study in New Zealand which result in the investigator breaking the blinding code at the time of the SAE or which result in hospitalisation or death.
- all serious adverse events occurring during the study worldwide which are considered related to the study medicine. Where there is a data safety monitoring board in place, serious adverse events occurring outside New Zealand may be reported quarterly.

All SAE reports must be signed by the Principal Investigator and include a comment on whether he/she considers there are any ethical issues relating to this study continuing due to this adverse event. It is assumed by signing the report, the Principal Investigator has undertaken to ensure that all New Zealand investigators are made aware of the event.

Amendments

All amendments to the study must be advised to the Committee prior to their implementation, except in the case where immediate implementation is required for reasons of safety. In such cases the Committee must be notified as soon as possible of the change.

Please quote the above ethics committee reference number in all correspondence.

The Principal Investigator is responsible for advising any other study sites of approvals and all other correspondence with the Ethics Committee.

It should be noted that Ethics Committee approval does not imply any resource commitment or administrative facilitation by any healthcare provider within whose facility the research is to be carried out. Where applicable, authority for this must be obtained separately from the appropriate manager within the organisation.

We wish you well with your study.

Yours sincerely

Alieke Dierckx
Upper South A Ethics Committee Administrator
Email: alieke_dierckx@moh.govt.nz

Appendix 2

Participant Information Sheet

# Participant Information Sheet	AUT UNIVERSITY TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

1. Study Background and Information

This research is part of a PhD study.

You are invited to take part in our study investigating the diagnosis of shoulder pain. The time required will be between 3-5 hours, spread over several visits, over a period ranging from 2 to 4 weeks depending upon your shoulder condition. Taking part is voluntary (your choice). If you decide not to take part, this will not affect your future health care. You are welcome to bring whanau/friend or support person with you to any of the appointments in this study.

What is the Study About?

The study is about the diagnosis of shoulder pain. During this part of the study, you will receive a clinical examination by the principal researcher, a shoulder X-Ray and ultrasound scan and two injections into the shoulder (on separate occasions). Depending upon your response to these injections, some patients may also see a Sports Medicine Physician who will refer you for a Magnetic Resonance Imaging (MRI) scan that will include an injection into the shoulder joint. Your participation in this part of the study may provide you with a faster and more accurate diagnosis of your shoulder problem than would normally occur, free of charge.

What is the Importance of the Study?

Shoulder pain is a very common, and disabling problem. Shoulder pain is very difficult to diagnose accurately for doctors and physiotherapists because of the complex anatomy involved, and the poor accuracy of many of the clinical tests currently used. This means that many people suffering from shoulder pain spend long periods of time with pain and disability before a diagnosis is made. This might delay treatment to give relief from symptoms, and a delay returning to work, sports or other daily activities.

What is the Purpose of the Study?

The purpose of this study is to see which of the findings from your clinical examination are most helpful for diagnosing your shoulder pain. We will find this out by comparing the clinical examination findings with what we find on various radiology procedures (X-Rays, Ultrasound Scans and Magnetic Resonance Arthrograms). We can also tell if a structure is the cause of pain by comparing your pain levels with certain clinical tests before and after an injection of local anaesthetic.

Who is Eligible to be Part of the Study?

You are eligible for this study if you are:

- You are 18 year of age or older
- Able to understand the English language
- Experiencing a first, or a new episode of shoulder pain that you have not yet seen anyone about or had any treatment for.
- Your pain is PRIMARILY in the shoulder region.
- You are available for, and consent to undergo all the procedures as outlined in this document.

Who Cannot be Part of the Study?

You are <u>not suitable</u> to be part of the study if:

- You do not consent to the procedures outlined in this document.
- You are pregnant or breastfeeding (due to the radiology procedures used)
- You have a fracture around the shoulder, or a dislocation of the shoulder or acromioclavicular joint.
- Your pain is being referred from your neck, or from somewhere else other than your shoulder.
- You have had previous surgery to your shoulder or neck.
- You have pins and needles, or numbness or significant unexplained weakness of your arm or hand.
- You are taking anticoagulant medications (blood thinners) eg, warfarin or heparin.
- You have had a previous adverse reaction to:
  - a local anaesthetic injection
  - contrast medium used during some X-Rays
- If you have any metal implants eg, stent, pacemaker, pins/screws, joint replacements (meaning you cannot have an MRI scan).
- If you <u>do not</u> wish your GP to be informed of your results.

NOTE: If any of these apply to you, please inform the researcher.

What is Involved in the Study?

We are aiming to recruit 200 participants into the study from the central Christchurch area between August 2009 and December 2010. If you give your consent to be part of this study, you will receive the following:

<u>**Visit 1:**</u>   Clinical Examination with Physiotherapist

<u>**Visit 2:**</u> X-Ray & Ultrasound scan - including injection of local anaesthetic into the subacromial bursa

<u>**Visit 3:**</u> Injection of local anaesthetic into the acromioclavicular joint under X-Ray guidance

If you experience a significant (more than 80%) reduction in your pain from the injections in Visit 2 or 3 (above), you will exit the study and return to your GP or physiotherapist for ongoing treatment/management of your shoulder pain.

If you do <u>NOT</u> experience 80% reduction in your pain following the injection of local anaesthetic in Visit 2 & 3 (above) you would proceed to:

<u>**Visit 4:**</u> Consultation with Sports Medicine Physician

<u>**Visit 5:**</u> Shoulder Arthrogram and MRI scan

<u>**Visit 6:**</u> Follow-up consultation with Sports Medicine Physician

Are These All Normal Procedures?

Yes, the tests and procedures being used in the study are normal. The tests are commonly and frequently used in the diagnosis and management of shoulder pain, and all the professionals involved are well trained and highly experienced in these procedures. The procedures are all safe, and adverse events relating to these procedures are extremely rare.

Where Would I have to go for these Appointments and Who Pays?

All these appointments will take place in the central Christchurch area. Exact locations can be found in Table 1 – below. Maps for all research locations are included in this document.

There is <u>no cost</u> to the participant for any of the procedures received as part of the research study. This includes the clinical examination, X-Rays, Ultrasound scan, injections, Sports Physician visits, and Magnetic Resonance Arthrogram.

This research does not involve any treatment and <u>does not cover any treatment costs</u>, orthopaedic specialist or surgical costs that may be received during or after your participation in the research study, including any treatment recommendations made by the Sports Physician.

What are the Benefits of Participating in this Study?

The main benefit is that you would receive a thorough set of diagnostic tests, designed to identify the source of your shoulder pain, and in a faster time-frame than is considered "normal". You will also receive these tests at no cost.

Your participation will also, benefit those suffering shoulder pain in the future as it is hoped this study may change the way shoulder pain is managed by GP's and Physiotherapists. The benefits and risks of specific procedures are described in more detail in the following sheets.

What are the Risks of Participating in the Research Study?

This study involves a number of medical investigation procedures. There are risks with any medical procedure. Although risks and side effects are rare – you should be aware of them. You will find a list of risks related to specific procedures used in this study throughout the Participant Information Sheet. Please ensure you read the risks associated with each procedure before signing the Participant Consent Form.

You should advise the principal researcher if you enrolled in any other clinical research studies.

What are the Alternatives to These Investigations?

All procedures used in this research study are currently considered the best procedures for determining the diagnosis of shoulder pain. Alternative investigations are sometimes used, eg, CT (computerised tomography) scans, bone scans and investigative surgery (arthroscopy), however these procedures involve higher doses of radiation, or are more invasive than the procedures chosen for this study. Your general practitioner will be able to provide more information about alterative procedures used in the diagnosis of shoulder pain should this be required.

Receiving Treatment for Your Shoulder Pain During the Study

We cannot withhold necessary treatment, however, results of the research may be compromised if treatment results in a change in your symptoms. We therefore ask that unless treatment is absolutely necessary, you refrain from treatment of your shoulder pain until the diagnostic studys in this research project are completed. It is expected the majority of these procedures will be completed within 1 week, with the others taking a maximum of three weeks.

Your Rights

You, or your representatives, i.e. relatives, guardians or, if necessary, legal representatives, will be given ample opportunity to enquire about details of the study and be allowed sufficient time to decide whether or not you wish to participate.

If you have any questions or concerns about your rights as a participant in this research study you can contact an independent health and disability advocate. This is a free service provided under the Health and Disability Commissioner Act.

Telephone: (NZ wide) 0800 555 050
Free Fax (NZ wide): 0800 2787 7678 (0800 2 SUPPORT)
Email (NZ wide): advocacy@hdc.org.nz

Table 1: Details of Personnel, Visits and Time Required for Procedures:

Diagnostic Shoulder Study - Procedures and Information

("Researcher" is Angela Cadogan (Physiotherapist))

Visit	Investigation	Procedure	Times	Total Time	Referral time-frames for Procedures	Procedure Conducted by:	Location/Time: (Maps of locations on pg 22)
1	Initial Assessment	Consent Forms & Questionnaires	30 mins	1 hour 15 mins	3-4 days from GP/physio referral.	Researcher	Physiosouth, Moorhouse Medical Centre, 3 Pilgrim Place. *(by appointment)*
		Clinical Examination	45 mins				
2	X-Ray, Ultrasound Scan & Bursal injection	Complete radiology forms and X-Ray	30 mins	1 hour & 15 mins	within 1 week from time of Clinical Examination	Researcher	Southern Cross Radiology, 129 Bealy Ave, Christchurch *(Wednesdays 12pm-3.30pm)*
		Clinical Reassessment	5 mins			Radiographer	
		Ultrasound Scan and injection of local anaesthetic into the bursa	30 mins			Sonographer + Radiologist	
		Clinical Reassessment	10 mins			Researcher	
3	Injection into the Acromioclavicular Joint	Administration and clinical assessment	20 mins	1 hour	1 week following Visit 2 (above)	Researcher	
		Injection under X-Ray guidance	30 mins			Radiologist	
		Clinical Reassesment	10 mins			Researcher	
4	Sports Physician Consultation	Sports Physician Consultation	20 mins	20 mins	within 1 week from time of Visit 3. (You will see only ONE of the Drs listed depending upon availability).	Dr Tony Page	Physiosouth - Profitness (Harvey Normal Centre), Moorhouse Ave. *(by appointment)*
						Dr Rob Campbell	SportsMed - 156 Bealey Ave, Christchurch. *(by appointment)*
						Dr Nat Anglem	Active Health, QEII Stadium, Burwood. *(by appointment)*
5	MR Arthrogram	Clinical Reassessment	5 mins	1 hour	within 1-2 weeks of Sports Physician Referral	Researcher	Hagley Radiology *(Thursdays 8am-11am ONLY)*
		Shoulder Arthrogram (injection)	15 mins			Radiologist	
		Clinical Reassessment	10 mins			Researcher	
		Magnetic Resonance Imaging Scan (MRI)	30 mins			MRI technicians	
6	Sports Physician	Follow-up Consultation	20 mins	20 mins	within 2-3 weeks of MRA procedure	same Dr you saw in Visit 4 (above)	*(by appointment)*
		MAXIMUM TOTAL TIME	**5 hours over 6 visits**		(approx 4-5 weeks)		

If you have a good response to injections in either visit 2 or 3 (above) and elect to undergo a corticosteroid injection, the corticosteroid injection will be performed the week following Visit 3. (you will <u>not</u> require visits 4-6).

Important Note:	Appointment times for X-Ray, ultrasound scans, injections and magnetic resonance arthrograms (MRA) (Visits 2,3 & 5) are fixed (**Wednesday afternoons 12pm-3.30pm, and Thursday mornings 8-11am**). No appointments will be available outside these times for Shoulder Research Study patients. If you are unable to attend at these times please inform the researcher as soon as possible.

Privacy & Confidentiality

- During the study, some personal information will be collected about you. The collection, storage and subsequent use of personal information will be strictly confidential and comply with the provisions of the privacy legislation.

- Hard copies of your clinical examination, imaging and test results will be kept in a locked filing cabinet (key access only) in the research office, and will be viewed only by the principal researcher, supervisors and the research assistant who will sign a confidentiality agreement.

- Once your personal information and results have been entered into the electronic data sheet, they will be allocated a numeric code thereby assuring you will not be personally identified in any way thereafter. The master list of names and associated numeric codes will be kept in a locked filing cabinet at the research office and will not be released to anyone.

- Any video and associated audio files will be kept on a password protected external hard drive on the principal researchers' computer. They will be kept for the purpose of checking the accuracy of the data collected at a later date and will not be viewed by anyone apart from the principal researcher and the two supervisors.

- You will be advised of the results of any procedures (X-Rays, Ultrasound Scan, Arthrogram and MRI), but should you wish to gain access to them, or any of your other personal information at other times, contact the principal researcher to arrange access.

- Your general practitioner (GP) will be sent a copy of all results/reports of your examination, injection and radiology procedures.

- After 10 years the hard copies of your information will be shredded.

- You will receive a copy of the final published report of the study.

- With your permission, data from this study may be used in future related studies, which have been given ethical approval from a Health & Disability Ethics Committee.

Termination of Participation in the Research Study

Your participation in the study will terminate under the following circumstances:

- You withdraw your consent to participate in the study (you may do this at any time and for any reason).

- You experience more than 80% relief of your symptoms as a result of administration of the local anaesthetic injection.

- You develop symptoms indicating your shoulder may not be the primary source of pain e.g., referred pain to the shoulder region from the neck/cervical spine, or upper limb neurological symptoms indicating nerve involvement.

- You develop unrelated medical conditions or complications during the study, or personal circumstances develop which render you unable to continue your participation in the study.

Ongoing Care Upon Termination of Your Participation in the Study

Upon termination of your participation in the study, you will be referred back to your primary care provider (GP or physiotherapist) for ongoing management and treatment of your shoulder problem. The researchers involved with this study are not involved in any aspect of treatment for your shoulder problem. This is the responsibility of your GP or physiotherapist. At the completion of your participation in the study, you will be offered copies of all your personal, medical, clinical examination, imaging and test results, including any video and audio files.

Adverse Reactions or Events

Adverse reactions or events related to this study may include (but are not limited to)**:**

- Aggravation of shoulder pain
- Adverse reaction or allergy to either the local anaesthetic or the contrast medium used in the imaging procedures.

Adverse events or reactions are unlikely, however should they occur:

- You should report any adverse reactions or events as soon as possible to the principal researcher (details below).
- If medical assessment or treatment is required, or compensation is required (see below) you will be directed to your GP for assessment and appropriate treatment.
- Any serious adverse reactions or events will be reported to the Ethics Committee overseeing this study and to the Centre for Adverse Reaction Monitoring (CARM).

Compensation

In the unlikely event of a physical injury as a result of your participation in this study, you <u>may</u> be covered by ACC under the Injury Prevention, Rehabilitation and Compensation Act. ACC cover is <u>not automatic</u> and your case will need to be assessed by ACC according to the provisions of the 2002 Injury Prevention Rehabilitation and Compensation Act. If your claim is accepted by ACC, you still may not receive compensation. This depends on a number of factors such as whether you are an earner or non-earner. ACC usually provides only partial reimbursement of costs and expenses and there may be no lump sum compensation payable. There is no cover for mental injury unless it is a result of physical injury. If you have ACC cover, generally this will affect your right to sue the investigators.If you have any questions about ACC, contact your nearest ACC office or the investigator.

Who Do I Contact for More Information?

For more information relating to <u>anything in this information sheet</u>, contact the principal researcher anytime:

Angela Cadogan – Physiotherapist (PhD Candidate)

phone 021 150 3731

emailacadogan@vodafone.co.nz

<u>**Appointment queries**</u>**:**

For appointment queries relating to you Clinical Examination time, please phone:

Carla Millar (Research Assistant) ph **022 6008692**

Angela Cadogan is a Christchurch-based physiotherapist who is completing her PhD in the Diagnosis of Shoulder Pain. She is enrolled through Auckland University of Technology (AUT) and is being supervised by Dr Wayne Hing (Associate Professor, AUT University), Dr Mark Laslett (specialist physiotherapist in Christchurch; Senior Research Fellow, AUT) and Dr Peter McNair (Professor, AUT University). Dr Mark Coates of Christchurch Radiology Group is assisting with the radiology procedures, and Dr Tony Page, Dr Nat Anglem and Dr Rob Campbell (Sports Physicians) are assisting with assessments and referrals for some radiology procedures.

2. Pain Relief (local anaesthetic) Injections

NOTE: It is not advisable to drive immediately following an injection of local anaesthetic (or corticosteroid).

Pre- and Post-Injection Clinical Assessment

Immediately prior to each injection procedure (subacromial bursa and acromioclavicular joint injections) you will be assessed by the researcher. Who will perform 3-4 clinical tests that previously provoked your pain during your initial clinical examination and will ask you to rate your pain on the 0-10 scale. This will be repeated again immediately after your injection procedure so we can determine how much pain relief you experienced from the injection. This will enable us to determine whether the injected structure is the likely source of your shoulder pain.

Subacromial Bursa Injections:

Procedure:

During the course of the diagnostic ultrasound scan (described above), you will receive an injection into your subacromial bursa (SAB) performed by a trained and experienced musculoskeletal radiologist. Ultrasound guidance is used so the radiologist can be sure the injection is placed accurately into the target structure.

You will be asked to lie down, the area will be sterilised, and the needle placed into the bursa and a small amount of local anaesthetic substance (xylocaine) will be injected into the bursa.

Injection into the
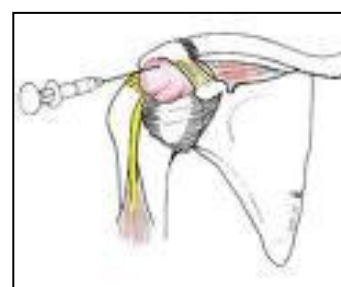
Acromioclavicular Joint Injections:

One week following the injection into your subacromial bursa, you will return and receive an injection of local anaesthetic into your acromioclavicular joint (ACJ) using "Fluoroscopy" to guide the needle.

A fluoroscope is a low-dose X-Ray machine that shows moving, real-time images. It is used in this study while the injection is being given, to ensure that the injection is placed accurately into the acromioclavicular joint

A
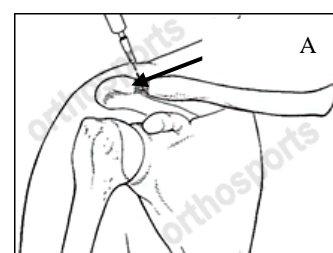
Injection into the acromioclavicular joint

Procedure

Firstly, a water soluble, iodinated (coloured) contrast medium is injected into the ACJ (along with the local anaesthetic - xylocaine). An X-Ray machine will be placed in front of your shoulder while the injection is being given, and images will be taken periodically to see the location of the needle, and the anaesthetic. Images look like a normal X-Ray. You will not feel anything during the fluoroscopy.

Note: This is the same procedure used for the injection of corticosteroid if you are involved in the Corticosteroid Injection Outcome Study.

Risks Associated with All Injection Procedures:

▪ Normal injection risks include some stinging, and infection (rare)
▪ Allergic reaction to the local anaesthetic
▪ Injections into the joint space (acromioclavicular joint) involve the risk of infection in the joint space.
▪ In rare cases people may experience a reaction to the contrast medium (see "Contrast Medium" page 17)

Benefits Associated with All Injection Procedures:

- You will receive this investigation at no cost
- You may experience rapid relief of your symptoms.
- Your response to the injection will assist making a clear diagnosis of your shoulder pain.

Your Response to the Injections:

If you experience an average of more than 80% relief of your symptoms following the injection of local anaesthetic into either the subacromial bursa or acromioclavicular joint, you will terminate your participation in the diagnostic study at this point.

If you do not experience greater than an average of 80% symptom relief following either of the injection procedures, you will proceed to the next step in the diagnostic study where you will see a Sports Medicine Physician and undergo a Shoulder Arthrogram and MRI scan.

What to Expect After the Injection Procedures:

The anaesthetic effect is only temporary, and will last between 4-8 hours. It is likely that if you experienced significant pain relief following the injection, that the pain will return after this time. This is normal.

It is also normal to experience some discomfort around the injection site or in your shoulder for approximately 24 hours following the procedure, however if this becomes progressively worse, or does not settle, see your own doctor for advice. If you experience any signs of infection including increasing redness around the injection site, increasing pain/tenderness of your shoulder, or feeling generally unwell, or if you are concerned about your response for any other reason, you should see your own doctor as soon as possible, and also notify the principal researcher in this study.

If you experience any reactions to the injection medication (see "Xylocaine" on page 18 and "Contrast Medium" page 17)or are concerned about your reaction for any other reason, you should also see your own doctor and notify the principal investigator.

3. Xylocaine (Lidocaine Hydrochloride) – Local Anaesthetic

Uses of xylocaine:
In the context of this study, this medication is used for anaesthetising (numbing) the structure that is injected. It works by blocking pain signals, allowing us to determine whether the injected structure is the source of your pain.

Benefits of xylocaine

- Fast acting pain relief may be experienced (lasting 2-4 hours)

Risks Associated with the Use of Xylocaine

- Usual injection risks apply (infection, stinging)
- Drug interactions: Beta-blockers (eg, atenolol) or digoxin. Tell the doctor about any medications you currently take.
- Allergy

Possible Side Effects of Xylocaine

All medicines may cause side effects, but many people have no, or minor, side effects. Check with your doctor if any of these most COMMON side effects persist or become bothersome:
- Temporary stinging/burning at the injection site
- Mild dizziness or drowsiness.

Seek medical attention under the following circumstances: Temporary loss of feeling, decreased muscle strength or dizziness may occur. If any of these effects persist or worsen, notify your doctor or pharmacist promptly.

Tell your doctor immediately if any of these unlikely but serious side effects occur: lightheadedness, nervousness, ringing in the ears, unusual pleasurable feelings/mood, confusion ,

dizziness, drowsiness/fatigue, nausea, decreased breathing, blurred/double vision, small pupils, vomiting, inability to urinate

Tell your doctor immediately if any of these highly unlikely but very serious side effects occur: twitching, restlessness, tremors, shivering, fever, slowed breathing, slowed/irregular heartbeat, seizures.

A serious **allergic reaction** to this drug is unlikely, but seek immediate medical attention if it occurs. Symptoms of a serious allergic reaction include:

- rash
- itching
- swelling
- severe dizziness
- trouble breathing.

If you notice other effects not listed above, contact your doctor or pharmacist.

Precautions:
Before receiving a lidocaine injection, tell your the doctor or researcher if you are allergic to it; or to certain preservatives (e.g., methylparaben); or to sulfites; or if you have any other allergies.

Before using this medication, tell your doctor or pharmacist your medical history, especially of: liver disease (e.g., chronic hepatitis, cirrhosis), severe shock due to the heart not pumping well (e.g., very low blood pressure and loss of consciousness), other heart problems (e.g., heart failure, previous heart attack), a certain abnormal heartbeat (heart block).

Maps of Research Locations

Physiosouth
Moorhouse Medical Centre *(2nd floor)*
3 Pilgrim Place, Waltham
- Initial research appointment occurs here.
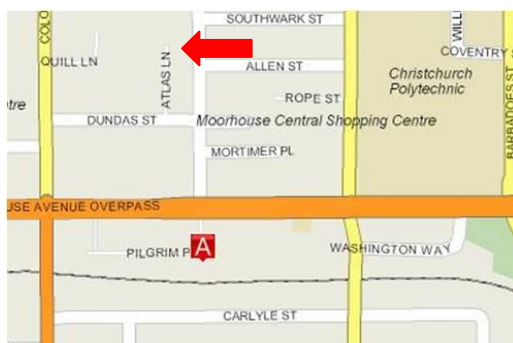
Southern Cross Radiology
129 Bealy Ave,Christchurch

Dr Tony Page
Physiosouth - Profitness
Harvey Normal Centre, Moorhouse Ave
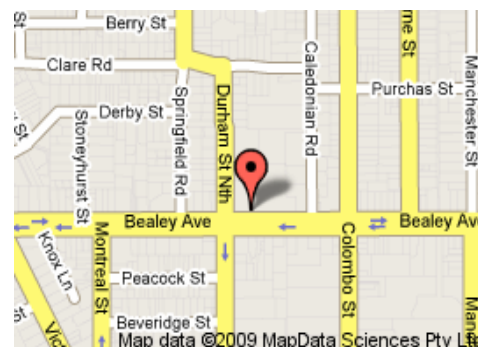
Dr Rob Campbell
SportsMed
156 Bealy Ave, Christchurch
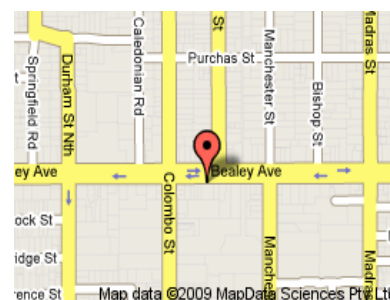
Appendix 3

Consent Form

Participant Consent Form	AUT UNIVERSITY TE WĀNANGA ARONUI O TĀMAKI MAKAU RAU

SHOULDER RESEARCH STUDY

Project title: *Diagnosis of Shoulder Pain at Primary Care Level*

Principal Researcher: **Angela Cadogan**

Project Supervisors: **Dr Wayne Hing, Dr Mark Laslett, Dr Peter McNair**

___

Consent for Participation in Research Study

(Please indicate your consent by placing a 'tick' (☑) in the boxes below)

☐ I have read and understood the information provided about this research project in the Information Sheet dated **22.7.09.** I understand the risks and benefits involved and have been given sufficient time to consider my participation in this study.

☐ I understand that my participation in this study is voluntary, and that I may withdraw from the study without reason at any time, without prejudice to legal and ethical rights, and without affecting my subsequent medical care.

☐ I (and my family/support person/whanau) have had an opportunity to ask questions and to have them answered.

☐ I understand that any information collected about me in this study will remain confidential and no information that can identify me will be used in reporting the results of this study.

☐ I understand the provisions for compensation in this study

☐ I agree to take part in this research.

Consent for Research Procedures

(Please indicate your consent by placing a 'tick' (☑) in the boxes below).

☐ Clinical Examination

☐ Shoulder X-Ray and Diagnostic Ultrasound Scan

☐ Injection of local anaesthetic into the subacromial bursa (under ultrasound guidance) and acromioclavicular joint (under X-Ray (fluoroscopic) guidance)

☐ Consultation and follow-up with Sports Physician

☐ Shoulder Arthrogram and Magnetic Resonance Imaging (MRI) scan

Consent for Release of Information

During the course of the study, some personal information will be collected. The collection, storage and subsequent use of personal information will be strictly confidential and comply with the provisions of the privacy legislation. In signing this consent form, you are consenting to the following:

- The collection of personal information that is relevant to the research study.

- An approved auditor appointed by either the ethics committee, or the regulatory authority or their approved representative, and approved by the Health Research Council ethics committee reviewing my relevant medical records for the sole purpose of checking the accuracy of the information recorded for the study.

- The viewing of the information by the following authorised research personnel only:
- The principal researcher
- The research assistant
- Research supervisors

- The release of your medical results from this study to your GP (General Medical Practitioner), and to the medical practitioner who referred you to this research study (GP or physiotherapist).

In addition, please indicate below whether you consent to the following (circle):

- I consent to the findings of any additional investigation or surgical procedures **Yes   No** being made available to the research personnel listed above.
- I consent to my interview and physical examination being video-taped, to be viewed for the checking of data accuracy.   **Yes   No**
- I wish to receive a copy of the videotape of the clinical examination   **Yes   No**
- I consent to the use of my data for future related studies, which have been given ethical approval from a Health & Disability Ethics Committee.   **Yes   No**
- I give permission to be contacted regarding the possibility of participating in further research studies into the treatment of shoulder pain at the conclusion of this study.   **Yes   No**

Results of the Study

- I wish to be notified of the results of this research (circle):   **Yes   No**

Address (if different from that already provided):

_____

_____

_____

Signatures

Participant:

Participant's name: ....………………………………. Date: __ / __ / __

Participant's signature: ....

Researcher:

Principal Researcher's signature:...Date: __ / __ / __
(Angela Cadogan)

Appendix 4

Published Manuscript:

Reliability of a New Hand-Held Dynamometer in Measuring Shoulder Range of Motion and Strength

Reprinted from:

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Williams, M. (2011). Reliability of a new hand-held dynamometer in measuring shoulder range of motion and strength. *Manual Therapy, 16*(1), 97-101. doi:10.1016/j.math.2010.05.005

with permission from Elsevier

Technical and measurement report

Reliability of a new hand-held dynamometer in measuring shoulder range of motion and strength

Angela Cadogan [a,b], Mark Laslett [a,b,c], Wayne Hing [a,b,*], Peter McNair [a,b], Maynard Williams [a,b]

[a] School of Rehabilitation and Occupation Studies, Faculty of Health & Environmental Sciences, AUT University, Private Bag 92006, Akoranga Drive, Northcote, Auckland 1142, New Zealand
[b] Health & Rehabilitation Research Institute, AUT University, Private Bag 92006, Auckland 1142, New Zealand
[c] Physiosouth, Christchurch, New Zealand

ABSTRACT

Acceptable reliability is a prerequisite for inclusion of physical examination tests in clinical examinations of the painful shoulder. The aim of this study was to establish the intraexaminer and interexaminer reliability of measures of shoulder range of motion (ROM) and muscle force using a new hand-held dynamometer with the ability to standardize overpressure force during passive ROM tests. Forty consecutive subjects with shoulder pain were recruited, and tests were performed by two physiotherapists. Tests included active ROM elevation, passive ROM glenohumeral abduction and external rotation and resisted abduction and external rotation. All tests demonstrated high levels of intraexaminer reliability (ICC 0.85–0.99; LOA 6–24° and 1.1–7.0 kg). Highest levels of interexaminer reliability reliability were observed for measures of active ROM flexion (ICC 0.88–0.95; LOA 14–22°). Passive ROM tests demonstrated 'moderate – substantial' interexaminer reliability (ICC 0.45–0.62; LOA 25–34°). The ICCs for resisted tests ranged from 0.68 to 0.84, and LOA ranged from 3.2 to 8.5 kg. Active ROM flexion demonstrated high levels of both intra- and interexaminer reliability. Measures of passive ROM and peak isometric force demonstrated acceptable levels of intraexaminer reliability.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Shoulder pain is a common complaint resulting in significant pain, functional disability and loss of quality of life (Turner-Bowker et al., 2003; MacDermid et al., 2004; Lin et al., 2005). The diagnosis of shoulder pain involves a clinical examination which typically consists of a variety of physical examination tests and associated measures including active and passive range of motion (ROM), and resisted muscle tests. The results of these tests including measures of active shoulder elevation and passive ROM abduction and external rotation are commonly used for diagnostic classification, and in the assessment of functional impairment (Constant and Murley, 1987; Davis, 1998; Harrington et al., 1998; MacDermid et al., 2007). Reliable measurements are required if these classifications are to be consistently applied.

Although few studies have directly compared reliability between active and passive ROM of the shoulder, more variability

has been reported in measures of passive ROM (ICC 0.26–0.90) versus active ROM (ICC 0.49–0.88) (Riddle et al., 1987; Hayes et al., 2001; Hoving et al., 2002; Terwee et al., 2005). A common explanation for this variability in measures of passive ROM is the inability of the examiner to standardize the amount of overpressure applied at the end range of motion (Boone et al., 1978; Gajdosik and Bohannon, 1987; Lea and Gerhardt, 1995; Hayes et al., 2001). A new hand-held dynamometer (HHD) (Industrial Research Ltd) has been developed that has the ability to simultaneously measure both angle and force. This feature enables the standardization of overpressure force at end range of motion. Whether this feature would reduce measurement variability and improve reliability during measures of passive ROM has not been tested to date.

Measures of muscle strength are used in the diagnostic process to assess muscle integrity and to determine the level of any strength deficits (Constant and Murley, 1987; Cyriax, 1982). Hand-held dynamometry has demonstrated higher sensitivity, and interexaminer reliability than manual muscle testing in identifying strength deficits of the rotator cuff (Ellenbecker, 1996; Leggin et al., 1996; Hayes et al., 2002; Tyler et al., 2005). Hand-held dynamometry therefore provides an advantage over manual muscle testing for the accurate clinical assessment of isometric muscle

* Corresponding author. Health & Rehabilitation Research Institute, AUT University, Private Bag 92006, Auckland 1142, New Zealand. Tel.: +64 9 9219999x7800; fax: +64 9 9219620.
 E-mail address: wayne.hing@aut.ac.nz (W. Hing).

strength. The reliability of the new HHD needs to be established for measurements of muscle strength before it can be used for this purpose in the clinical setting.

Thus, the aim of this study was to establish the intra- and interexaminer reliability of a new HHD in measuring ROM and isometric muscle strength of the symptomatic shoulder. Whether reliability of measures of passive ROM could be improved by standardizing the amount of overpressure force applied at end range of motion was specifically investigated.

2. Methods

2.1. Subjects

Forty consecutive subjects with shoulder pain were recruited from local physiotherapy practices. Subjects were included if they were over 18 years of age and were currently experiencing shoulder pain. Subjects were excluded if they had pain referred from a source other than the shoulder, fractures or dislocations around the shoulder joint, or were suffering known systemic inflammatory disease. Ethical approval was gained from the Ministry of Health Ethics Committee.

2.2. Hand-held dynamometer

The Industrial Research Ltd hand-held dynamometer (HHD) has the ability to simultaneously measure angle (°) and force (kg) (Fig. 1a). This feature enables standardization of load applied at end of range of motion during passive movement testing by selecting a force level at which an audible alarm is produced. Force is recorded from an in-built force transducer attached to the HHD, and peak force (kg) is displayed. The HHD is gravity dependent, and indicates range of motion on a 360° scale with reference to the vertical plane. The HHD records absolute range of motion calibrated from a zero position, and final range of motion. The relative range is displayed on the unit being calculated by subtracting the initial starting range from the final range of motion. Each examiner used a separate HHD, and both were calibrated on the first day of data collection to ±1° and to within ±0.1 kg of force.

2.3. Procedures

Prior to the study, the examiners underwent 4 sessions of familiarization training with the HHD. Seven physical examination tests were performed. Range of motion was measured during active ROM elevation (through flexion), passive ROM glenohumeral abduction and external rotation (performed at 0° of abduction). Peak isometric force (kg) was measured during resisted abduction and external rotation on both the affected and unaffected sides.

The physical examination tests were performed on the same day by two experienced physiotherapists (19 and 38 years experience). The examiners were blinded to each others results. Examiner sequence and the order of tests were randomized using a random sequence generator for each subject and each examiner.

2.3.1. Range of motion tests

For active ROM elevation through flexion, each subject stood with their back against the wall to prevent compensatory movement of the trunk. The HHD was aligned along the long axis of the humerus and three trials were performed. Each trial was followed by approximately 30 s of rest. For tests of passive ROM glenohumeral abduction, subjects were positioned as in Fig. 1b. The examiner applied firm downward pressure over the acromion while the subjects' arm was guided into abduction in the scapula plane. To compensate for variation in subject limb mass when raising the arm against gravity, a standardized force equivalent to 6% of the subjects' body mass was programmed into the HHD and when this force level was reached, the sound of the audible alarm was used as the criteria for end range. The 6% body mass level was determined following pilot testing on a sample of patients with shoulder pain. This force level was consistently found to be required in order to reach end range of motion without excessive discomfort to the patient.. Measurement of shoulder external rotation is shown in Fig. 1c. A load was required that would overcome the mass of the arm to allow end range of motion to be achieved, without causing excessive discomfort to the subject. Average upper limb mass is approximately 3–4% of total body mass (Clarys and Marfell-Jones, 1986), and based upon results of pilot testing a standardized load of 3 kg was selected as it appeared to consistently fulfill end range and comfort criteria for all subjects.

2.3.2. Resisted muscle tests

Resisted abduction was performed with the subject sitting with the arm in 10° of abduction in the plane of the scapula. The HHD was placed immediately proximal to the lateral epicondyle of the humerus. Resisted external rotation was performed with the subject sitting with the forearm in neutral, the arm by the side, exerting a slight adduction force against the examiner's hand while simultaneously exerting an external rotation force against the HHD held against the distal forearm. Three maximal isometric contractions — were performed, the duration of which was approximately 6–7 s and 30 s rest was given between trials. Subjects were

Fig. 1. A – Hand-held dynamometer (Industrial Research Ltd, Christchurch, New Zealand); B – position for testing passive ROM glenohumeral abduction; C- position for testing passive ROM external rotation.

Table 1
Summary of subject characteristics.

Subject characteristics		Number	%
Gender			
Male		23	58
Female		17	42
Affected side			
Dominant		24	60
Non-dominant		11	28
Bilateral		5	13
		Mean	Range
Age (years)		49	18–77
Height (cm)		171	157–189
Weight (kg)		80	53–102
Duration of symptoms (months)		48	<1–325
Pain severity in previous 24 h (11 point VAS)		3.6	0–7

instructed to hold the contraction against maximal examiner pressure and peak isometric muscle force was recorded.

2.4. Statistical methods

Relative reliability was assessed using single-measure intraclass correlation coefficients ($ICC_{2,1}$) and associated 95% confidence intervals (CI) (two-way random effects model -absolute agreement). For intraexaminer reliability, data from the first trial was compared with data from the second trial, and the mean of three trials for both examiners. For interexaminer reliability, data from a single trial (first trial (ROM) or the peak force trial (resisted tests)), and the single trial compared with the mean of three trials were used for the analysis. A one-way ANOVA was used to ascertain any differences between trials (intraexaminer reliability) and between examiners (interexaminer reliability) with the level of significance set at $p = 0.05$.

Absolute reliability was determined by calculating the mean difference between measures and the associated 95% CI for the mean difference, as well as limits of agreement (LOA) according to the Bland and Altman method of assessing agreement (mean difference between examiners $+/-1.96$ SD_{diff}) (Bland and Altman, 1986).

Reliability values were interpreted according to the guidelines of Landis and Koch (1977); 0.00–0.20 slight; 0.21–0.40 fair; 0.41–0.60

moderate; 0.61–0.80 substantial; 0.81–1.00 almost perfect (Landis and Koch, 1977).

3. Results

Demographic characteristics of the subjects are provided in Table 1. Twenty one subjects were examined first by examiner 1 and 19 by examiner 2.

3.1. Intraexaminer reliability

Mean differences between Trial 1 and 2, and between Trial 1 and the mean of three trials for both examiners, LOA and ICC values are presented in Table 2. For measures of ROM, ICCs ranged from 0.85 (passive ROM abduction) to 0.99 (active ROM elevation). Limits of agreement ranged from $+6°$ (active ROM elevation) to $+24°$ (passive ROM abduction). For resisted tests ICCs ranged from 0.91 to 0.99, and LOA ranged from $+1.1$ kg (resisted external rotation -unaffected side) to $+7.0$ kg (resisted abduction $-$ affected side). The results of comparisons between Trial 1 and the mean of three trials showed consistently higher levels of reliability and agreement than comparisons between single trials (trial 1 and 2) for all tests. A significant difference between the mean of the trials ($p < 0.05$) was identified for resisted external rotation for examiner 2 (0.1–0.2 kg) (Table 2).

3.2. Interexaminer reliability

Mean differences between examiners, LOA and ICC values for interexaminer reliability are presented in Table 3. For measures of ROM, ICCs ranged from 0.45 (passive ROM abduction) to 0.95 (active ROM elevation $-$ mean of three trials). Widest 95% CI were observed for measures of passive ROM. Limits of agreement ranged from 14° (active ROM elevation $-$ mean of 3 trials) to 34° (passive ROM abduction $-$ 1st trial). For resisted tests, ICC values ranged from 0.68 to 0.84, with the widest 95% CI recorded for resisted external rotation (unaffected side). Limits of agreement ranged from 3.2 kg (resisted external rotation $-$ mean of 3 trials) to 8.5 kg (resisted abduction $-$unaffected side).

Table 2
Intraexaminer reliability: mean differences between trials and limits of agreement (LOA) with 95% confidence intervals (CI) for mean difference, intraclass correlation coefficients (ICC) and associated 95% confidence intervals for active and passive ROM and peak isometric force.

Physical examination test		Examiner 1			Examiner 2		
		Mean (Range)	Mean Difference Between Trials +/− LOA	ICC (95% CI)	Mean (Range)	Mean Difference Between Trials +/− LOA	ICC (95% CI)
Range of motion (°)							
Active elevation	Single[a]	150 (84–181)	0.5 ± 19 (−2.5, 3.5)	0.92 (0.88, 0.96)	152 (94–182)	0.8 ± 9 (−2.1, 0.6)	0.98 (0.97, 0.99)
	Mean of 3 trials[b]	150 (88–180)	0.7 ± 14 (−1.6, 3.0)	0.96 (0.92, 0.98)	153 (95–180)	0.2 ± 6 (−1.0, 0.7)	0.99 (0.99, 1.0)
Passive abduction	Single	76 (35–126)	0.4 ± 24 (−4.2, 3.5)	0.85 (0.76, 0.91)	86 (44–128)	0.8 ± 14 (−1.4, 2.9)	0.91 (0.85, 0.95)
	Mean of 3 trials	77 (36–116)	0.5 ± 13 (−1.6, 2.6)	0.94 (0.88, 0.97)	86 (45–119)	0.5 ± 9 (−1.0, 1.9)	0.96 (0.93, 0.98)
Passive external rotation	Single	46 (8–126)	1.0 ± 12 (−2.9, 0.9)	0.89 (0.83, 0.94)	61 (9–117)	2.0 ± 13 (−4.1, 0.2)	0.95 (0.92, 0.97)
	Mean of 3 trials	45 (8–106)	1.0 ± 10 (−2.6, 0.6)	0.96 (0.93, 0.98)	61 (10–112)	0.8 ± 8 (−2.1, 0.5)	0.98 (0.97, 0.99)
Resisted muscle tests (kg)							
Resisted abduction (AS)[c]	Single	20.0 (4.9–37.2)	0.8 ± 7 (−2.0, 0.3)	0.91 (0.85, 0.95)	17.0 (5.5–29.9)	0.1 ± 3.6 (−0.5, 0.7)	0.95 (0.92, 0.98)
	Mean of 3 trials	19.6 (5.0–33.1)	0.4 ± 4.1 (−1.0, 0.3)	0.96 (0.93, 0.98)	17.3 (6.4–30.7)	0.3 ± 2.2 (−0.1, 0.6)	0.98 (0.97, 0.99)
Resisted abduction (US)[d]	Single	21.0 (8.5–34.0)	0.4 ± 4.5 (−1.2, 0.3)	0.95 (0.91, 0.97)	19.0 (7.7–34.3)	0.2 ± 3.8 (−0.4, 0.8)	0.95 (0.92, 0.97)
	Mean of 3 trials	21.3 (9.0–34.2)	0.2 ± 3.0 (−0.7, 0.4)	0.98 (0.96, 0.99)	19.3 (7.9–33.5)	0.2 ± 2.7 (−0.2, 0.7)	0.98 (0.96, 0.99)
Resisted external rotation (AS)	Single	11.0 (4.6–19.7)	0.1 ± 1.7 (−0.2, 0.3)	0.96 (0.94, 0.98)	13.0 (4.7–20.4)	0.6 ± 3.2 (0.1, 1.1)	0.91 (0.85, 0.95)*
	Mean of 3 trials	11.3 (5.1–19.7)	0.0 ± 1.1 (−0.2, 0.1)	0.99 (0.97, 0.99)	12.8 (5.2–20.4)	0.5 ± 2.0 (0.1, 0.8)	0.96 (0.90, 0.98)*
Resisted external rotation (US)	Single	12.0 (5.3–18.6)	0.3 ± 1.9 (−0.6, 0.0)	0.94 (0.90, 0.97)	14.0 (5.6–22.3)	0.7 ± 3.0 (0.2, 1.2)	0.93 (0.87, 0.96)*
	Mean of 3 trials	11.9 (5.9–17.9)	0.2 ± 1.4 (−0.4, 0.0)	0.98 (0.95, 0.99)	13.9 (5.8–20.4)	0.4 ± 1.7 (0.1, 0.7)	0.97 (0.92, 0.98)*

AS, affected side; US, unaffected side.
*One-way ANOVA significant difference between trials ($p < 0.05$).
[a] Trial 1 compared to Trial 2.
[b] Trial 1 compared to mean of 3 trials.

Table 3
Interexaminer reliability: mean and range of values, mean differences between examiners and limits of agreement (LOA) with 95% confidence intervals (CI) for the mean difference, intraclass correlation coefficient (ICC) values with associated 95% confidence intervals for range of motion and peak isometric force.

Test		Mean values (range)	Mean Diff ± LOA (95% CI for Mean Diff)	ICC (95% CI)
Range of motion (°)				
Active elevation	1st trial	151 (84–182)	3.0 ± 22 (−0.5, 6.6)	0.88 (0.78, 0.93)
	Mean of 3 trials	152 (88–180)	2.1 ± 14 (−0.1, 4.4)	0.95 (0.90, 0.97)
Passive abduction	1st trial	81 (39–126)	9.0 ± 34 (3.4, 14.6)	0.45 (0.15, 0.67)*
	Mean of 3 trials	73 (36–119)	8.9 ± 30 (4.0, 13.9)	0.49 (0.17, 0.71)*
Passive external rotation	1st trial	54 (9–126)	15.3 ± 31 (10.2, 20.4)	0.58 (0.04, 0.81)*
	Mean of 3 trials	53 (8–112)	15.5 ± 30 (11.4, 19.5)	0.62 (−0.04, 0.85)*
Resisted muscle tests (kg)				
Abduction (AS)	Peak force trial	19.8 (5.4–37.2)	−2.9 ± 7.3 (−4.2, −1.8)	0.81 (0.42, 0.92)*
	Mean of 3 trials	18.5 (5.0–33.1)	−2.4 ± 6.3 (−3.4, −1.3)	0.84 (0.54, 0.93)*
Abduction (US)	Peak force trial	21.5 (8.4–36.7)	−2.5 ± 8.5 (−4.0, −1.0)	0.77 (0.49, 0.89)*
	Mean of 3 trials	20.3 (8.0–34.1)	−2.5 ± 7.6 (−3.8, −1.2)	0.77 (0.46, 0.89)*
External rotation (AS)	Peak force trial	12.7 (5.5–21.1)	1.8 ± 4.4 (1.1, 2.6)	0.69 (0.23, 0.87)*
	Mean of 3 trials	12.1 (5.1–20.4)	1.5 ± 4.0 (0.9, 2.2)	0.74 (0.36, 0.89)*
External rotation (US)	Peak force trial	13.6 (6.1–22.3)	2.4 ± 3.6 (1.8, 3.1)	0.68 (−0.05, 0.89)*
	Mean of 3 trials	12.9 (5.8–20.4)	2.2 ± 3.2 (1.7, 2.8)	0.70 (−0.04, 0.90)*

AS, affected side; US, unaffected side.
*One-way ANOVA significant difference between examiners ($p < 0.05$).

It should be noted that zero was not within the 95% CI for the mean difference between examiners for any of the passive ROM or resisted muscle tests suggesting the presence of systematic bias. One-way ANOVA results also indicated significant differences between examiners for all passive ROM (1–8°) and resisted muscle tests (0.6–1.3 kg) (Table 3).

4. Discussion

4.1. Intraexaminer reliability

All measures of ROM and peak isometric force used in this study demonstrated clinically acceptable levels of intraexaminer reliability (ICC 0.85–0.99). The results were higher than previous intraexaminer reliability results for measures of active ROM elevation through flexion (ICC 0.49–0.88) (Hayes et al., 2001; Hoving et al., 2002; Terwee et al., 2005), passive ROM abduction (ICC 0.58–0.67) and passive ROM external rotation (ICC 0.60–0.73) (Hayes et al., 2001; Terwee et al., 2005). However previous studies compared measurements on two separate occasions. In the present study, consecutive trials were conducted on one occasion to assess the number of trials required for satisfactory reliability (ICC >0.80). Intraexaminer reliability results for peak isometric force measures were similar to those of Hayes et al. (2002) and Leggin et al. (1996) for resisted abduction (ICC 0.84–0.96) and external rotation (0.89–0.95). Differences between trials for examiner 2 during resisted external rotation may be due to increasing familiarity of the subject with the test with subsequent trials, or alterations in examiner technique following performance of the initial trial. In summary, for the purposes of assessing diagnostic criteria and physical impairments, the tests used in this study demonstrated acceptable intraexaminer reliability within a single clinical session.

4.2. Interexaminer reliability

4.2.1. Range of motion

Measures of active ROM elevation through flexion using the new HHD reached ICC values in excess of 0.80, and 95% limits of agreement between 14 and 22°. These results were higher than previous results for measures of active ROM elevation through flexion (ICC 0.65) (Hoving et al., 2002), and LOA (27–36°) (Triffitt et al., 1999).

Despite the ability to standardize overpressure force in the current study, interexaminer reliability for measures of passive ROM of the shoulder were lower than previously reported (ICC 0.64–0.92, SEM 7.5–14°) (Boone et al., 1978; Riddle et al., 1987; MacDermid et al., 1999; Hayes et al., 2001). There are several possible explanations for this finding.

This is a new device, and the amount of familiarization time required may have been underestimated. More experience with the HHD may reduce measurement variability resulting from subtle changes in planar angulation of the HHD. Factors relating to standardization of the test procedures including subject positioning and instructions to subjects may have also affected reliability. However care was taken to ensure that errors associated with these factors were addressed. The systematic error present in the results of passive ROM tests suggests an examiner source of error was present despite the care being taken to standardize procedures prior to the study.

A force relative to body weight (6%) was used as our criteria for end ROM during passive glenohumeral abduction to overcome the weight of the limb as it was lifted against gravity. Measures of passive external rotation did not require moving the limb against gravity, and an absolute force level of 3 kg was selected as the criteria for end range. No relationship between body weight and end ROM was identified during pilot testing, and this amount of load appeared to consistently achieve end range of motion for all subjects. This is not surprising given the average mass of the arm in this study was approximately 2.4–3.2 kg (3–4% of total body mass) (Clarys and Marfell-Jones, 1986) and the 3 kg load was likely to move the arm to its final end range of motion.

4.2.2. Resisted muscle tests

The results of the present study fall within the range of previously reported interexaminer reliability results for peak isometric muscle force during resisted abduction (ICC 0.79–0.92) (Leggin et al., 1996; Hayes et al., 2002) (Leggin et al., 1996; Hayes et al., 2002) and slightly lower than previously reported for resisted external rotation (ICC 0.82–0.94) (Leggin et al., 1996; Hayes et al., 2002; Hayes and Petersen, 2003). Peak isometric force measures during resisted abduction (affected side) in the present study demonstrated high levels of reliability, however LOA indicate that 95% of measurements between examiners would lie within a range either 7.3 kg higher or lower than the other examiner, and the lower CI for the ICC values indicated only 'moderate' reliability. This level

of variability suggests measurements between examiners should be interpreted with caution.

Other peak isometric force measures also demonstrated wide LOA and while the ICC values were slightly higher for the mean of three trials compared with single trials during resisted muscle tests, lower 95% CI were poor (-0.5–0.36) indicating caution should also be used when interpreting these results. Confidence intervals for the mean difference between examiner measures (zero not contained in the confidence intervals) and ANOVA results also indicate systematic and significant differences between examiners for these measures (Table 3). This may be due to the known limitations of the use of HHD, including examiner strength (Wadsworth et al., 1992), systematic differences in test procedure between examiners or a change in subject symptom severity with repeated testing. We recommend caution be exercised when interpreting peak isometric muscle force measures during both resisted abduction and external rotation.

5. Conclusion

Measures of active ROM elevation (flexion) obtained using the HHD were reliable within- and between examiners during one clinical session. Measures of peak isometric force during resisted abduction and external rotation, and measures of passive ROM abduction and external rotation using the HHD to standardize overpressure force also demonstrated clinically acceptable levels of intraexaminer reliability.

References

Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.

Boone DC, Azen SP, Lin CM, Spence C, Baron C, Lee L. Reliability of goniometric measurements. Physical Therapy 1978;58:1355–60.

Clarys JP, Marfell-Jones MJ. Anatomical segmentation in humans and the prediction of segmental masses from intra-segmental anthropometry. Human Biology 1986;58(5):771–82.

Constant CR, Murley AHG. A clinical method of functional assessment of the shoulder. Clinical Orthopaedics and Related Research 1987;214:160–4.

Cyriax JH. Textbook of orthopaedic medicine. In: Diagnosis of soft tissue lesions. 8th ed., vol. 1. Minneapolis: Balliere Tindall; 1982. p. 1–50.

Davis TRC. Diagnostic criteria for upper limb disorders in epidemiological studies. Journal of Hand Surgery (British and European) 1998;23B(5):567–9.

Ellenbecker TS. Muscular strength relationship between normal grade manual muscle testing and isokinetic measurement of shoulder internal and external rotators. Isokinetics and Exercise Science 1996;6:51–6.

Gajdosik RL, Bohannon RW. Clinical measurement of range of motion; review of goniometry emphasizing reliability and validity. Physical Therapy 1987;67:1867–72.

Harrington JM, Carter JT, Birrell L, Gompertz D. Surveillance case definitions for work related upper limb pain syndromes. Occupational and Environmental Medicine 1998;55:264–71.

Hayes K, Walton J, Szomor ZL, Murrell GAC. Reliability of 3 methods for assessing shoulder strength. Journal of Shoulder and Elbow Surgery 2002;11(1):33–9.

Hayes K, Walton JR, Szomor ZR, Murrell GA. Reliability of five methods for assessing shoulder range of motion. Australian Journal of Physiotherapy 2001;47(4):289–94.

Hayes KW, Petersen CM. Reliability of classifications derived from Cyriax's resisted testing in subjects with painful shoulders and knees. Journal of Orthopaedic and Sports Physical Therapy 2003;33(5):235–46.

Hoving JL, Buchbinder R, Green S, Forbes A, Bellamy N, Brand C, et al. How reliable do rheumatologists measure shoulder movement? Annals of Rheumatic Diseases 2002;61:612–6.

Landis JR, Koch GC. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

Lea R, Gerhardt J. Range of motion measurements. Journal of Bone & Joint Surgery – American Volume 1995;77-A(5):784–98.

Leggin BG, Neuman RM, Iannotti JP, Williams GR, Thompson EC. Intrarater and interrater reliability of three isometric dynamometers in assessing shoulder strength. Journal of Shoulder and Elbow Surgery 1996;5(1):18–24.

Lin J, Hanten WP, Olson SL, Roddey TS, Soto-quijano DA, Lim HK, et al. Functional activity characteristics of individuals with shoulder dysfunctions. Journal of Electromyography & Kinesiology 2005;15(6):576–86.

MacDermid JC, Chesworth BM, Patterson S, Roth JH. Intratester and intertester reliability of goniometric measurement of passive lateral shoulder rotation. Journal of Hand Therapy 1999;12:187–92.

MacDermid JC, Ghobrial M, Quirion KB, St-Amour M, Tsui T, Humphreys D, et al. Validation of a new test that assesses functional performance of the upper extremity and neck (FIT-HaNSA) in patients with shoulder pathology. BMC Musculoskeletal Disorders 2007;8:42.

MacDermid JC, Ramos J, Drosdowech D, Faber K, Patterson S. The impact of rotator cuff pathology on isometric and isokinetic strength, function, and quality of life. Journal of Shoulder & Elbow Surgery 2004;13(6):593–8.

Riddle DL, Rothstein JM, Lamb RL. Goniometric reliability in a clinical setting. Shoulder measurements. Physical Therapy 1987;67(5):668–73.

Terwee CB, de Winter AF, Scholten JP, Jans MP, Deville W, van Schaardenburg D, et al. Interobserver reproducibility of the visual estimation of range of motion of the shoulder. Archives of Physical Medicine and Rehabilitation 2005;86(7):1356–61.

Triffitt PD, Wildin C, Hajioff D. The reproducibility of measurement of shoulder movement. Acta Orthopedica Scandinavica 1999;70:322–4.

Turner-Bowker DM, Bayliss MS, Ware JEJ. Usefulness of the SF-8 health survey for comparing the impact of migraine and other conditions. Quality of Life Research 2003;12:1003–12.

Tyler TF, Nahow RC, , et alTyler TF, Nahow RC, Nicholas SJ, McHugh MP. Quantifying shoulder rotation weakness in patients with shoulder impingement. Journal of Shoulder & Elbow Surgery 2005;14(6):570–4.

Wadsworth CT, Nielsen DH, Corcoran DS, Phillips CE, Sannes TL. Interrater reliability of hand-held dynamometry: effects of rater gender, weight, and grip strength. Journal of Orthopaedic & Sports Physical Therapy 1992;16(2):74–81.

Appendix 5

Published Manuscript:

Interexaminer reliability of orthopaedic special tests used in the assessment of shoulder pain

Reprinted from:

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Williams, M. (2011). Interexaminer reliability of orthopaedic special tests used in the assessment of shoulder pain. *Manual Therapy, 16*(2), 131-135. doi:10.1016/j.math.2010.07.009

with permission from Elsevier

Original article

Interexaminer reliability of orthopaedic special tests used in the assessment of shoulder pain

Angela Cadogan [a,*], Mark Laslett [b], Wayne Hing [a], Peter McNair [a], Maynard Williams [c]

[a] School of Rehabilitation and Occupation Studies, Health and Rehabilitation Research Institute, AUT University, Auckland, New Zealand
[b] Physiosouth, Christchurch, New Zealand
[c] School of Rehabilitation and Occupation Studies, Faculty of Health & Environmental Sciences, AUT University, Auckland, New Zealand

ARTICLE INFO

ABSTRACT

Orthopaedic special tests (OST) are commonly used in the assessment of the painful shoulder to assist to rule-in or rule-out specific pathology. A small number of tests with high levels of diagnostic accuracy have been identified but interexaminer reliability data is variable or lacking. The aim of this study was to determine the interexaminer reliability of a group of OST with demonstrated diagnostic accuracy at primary care level. Forty consecutive subjects with shoulder pain were recruited. Six tests were performed by two examiners (physiotherapists) on the same day. Tests included the active compression test, Hawkins−Kennedy test, drop-arm test, crank test, Kim test and belly-press test. 'Fair' reliability (kappa 0.36−0.38) was observed for the active compression test (labral pathology), Hawkins−Kennedy test and crank test. Prevalence of positive agreements was low for the active compression test (acromioclavicular joint), drop-arm test, Kim test and belly-press test. Prevalence and bias adjusted kappa (PABAK) values indicated 'substantial' reliability (0.65−0.78) for these tests. The active compression test (acromioclavicular joint), belly-press tests (observation and weakness), Kim test and drop-arm test demonstrate acceptable levels of interexaminer reliability in a group of patients with sub-acute and chronic shoulder conditions.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The diagnosis of shoulder pain presents a significant challenge to the primary care clinician due to complex regional anatomy and the frequent coexistence of multiple pathologies. In order to reach a differential diagnosis of shoulder pain, clinicians commonly use orthopaedic special tests (OST) during the physical examination to assist with ruling-in, or ruling-out specific pathology (Cyriax, 1982). The results of these tests frequently form the basis for diagnostic and intervention decisions.

For a test to be clinically valid, acceptable levels of diagnostic accuracy and interexaminer reliability must be demonstrated (Fritz and Wainner, 2001). Of the OST used in the clinical examination of the painful shoulder, only a small number have demonstrated sufficient diagnostic accuracy to be of clinical use. The Hawkins−Kennedy test has been advocated as a useful screening test for 'impingement' lesions, and the belly-press test (subscapularis

muscle tear), and active compression test (acromioclavicular joint) are suggested to be specific for their respective pathologies (Hegedus et al., 2008). Anatomical validity has also been established for the Hawkins−Kennedy test and the active compression test (Green et al., 2008). The belly-press test has demonstrated validity for primary activation of the subscapularis muscle (Tokish et al., 2003). The crank test (superior glenoid labrum tears) (Liu et al., 1996; Mimori et al., 1999) and the Kim et al. (2005) test (posterior glenoid labrum tears) have demonstrated high levels of diagnostic accuracy for glenoid labral lesions, and the drop-arm test for a complete tear of supraspinatus (Codman, 1934; Murrell and Walton, 2001; Park et al., 2005).

While previous authors provide valuable analyses of the diagnostic accuracy and anatomical validity of OST, reliability data on many of these tests is lacking, and where available, demonstrates widely variable interexaminer reliability (Table 1). In many of these studies, confidence intervals (CI) and raw agreement statistics (percent agreement) are not reported, and only one study was identified in which subjects were recruited from primary care (Johansson and Ivarson, 2009). No studies investigating interexaminer reliability of the belly-press test were found.

* Corresponding author. Health & Rehabilitation Research Institute, AUT University, Private Bag 92006, Auckland 1142, New Zealand. Tel.: +64 21 1503731; fax: +64 3 3770614.
E-mail address: acadogan@vodafone.co.nz (A. Cadogan).

Table 1
Summary of previous interexaminer reliability studies of orthopaedic special tests for the shoulder.

Test	Author	Subject numbers	Examiners	Interexaminer reliability	
				Kappa (95% confidence intervals)	% agreement
Impingement Tests:					
Hawkins–Kennedy test	Johansson and Ivarson (2009)	33	Physiotherapists	0.91[b]	NR[a]
	Nanda et al. (2008)	63	Orthopaedic consultant and registrar	0.55[b]	95%
	Ostor et al. (2004)	136	Consultant, specialist registrar & nurse	0.18–0.43[b]	NR
	Razmjou et al. (2004)	136	Orthopaedic surgeon & physical therapist	0.29 (0.18, 0.40)	60%
	Norregaard et al. (2002)	86	Orthopaedic surgeon & rheumatologist	0.07–0.40[b]	
Glenoid Labrum Tests:					
Kim test	Kim et al. (2005)	172	Orthopaedic surgeons	0.91[b]	NR
Crank test	Walsworth et al. (2008)	55	Orthopaedic surgeons & physical therapist	0.20 (−0.05, 0.46)	60%
Active compression test	Walsworth et al. (2008)	55	Orthopaedic surgeons & physical therapist	0.24 (−0.02, 0.50)	60%
Rotator Cuff Integrity Tests:					
Drop-arm test	Nanda et al. (2008)	63	Orthopaedic consultant and registrar	0.35[b]	77%
	Ostor et al. (2004)	136	Consultant, specialist registrar & nurse	0.28–0.66[b]	NR

[a] NR – not reported.
[b] Confidence intervals not available.

The aim of this study was to determine the interexaminer reliability of two experienced physiotherapists in determining the results of a selection of OST with known diagnostic accuracy used in the assessment of shoulder pain in patients recruited from primary care. The results will inform the content of a clinical examination to be used as index tests in future diagnostic studies, and serve as a guide for the clinician to the selection of evidence-based OST for use in clinical examination of the painful shoulder.

2. Methods

2.1. Subjects

Consecutive subjects were recruited through physiotherapy practices in Christchurch, New Zealand. Subjects were included in the study if they were over 18 years of age and currently experiencing shoulder pain. Subjects were excluded where pain was referred from a source other than the shoulder, or if there was a history of fracture or dislocation to the shoulder. The study was approved by the Ministry of Health Ethics Committee.

2.2. Procedures

OST were performed by two experienced examiners (19 and 38 years experience) on the same day. In order to prevent the occurrence of systematic differences between the examiners due to repeated testing and changes in subjects' symptom response following the first assessment, the sequence of the examiners was

randomly allocated. The order of tests was also randomized using a random sequence generator for each subject and each examiner. Prior to the study, the examiners underwent four training sessions to standardize test procedures and to familiarize themselves with the use of the hand-held dynamometer used to measure peak isometric muscle force during the belly-press test.

The OST were selected according to diagnostic accuracy values and those identified as being of clinical value by Hegedus et al., (2008). The associated criteria for a positive test result are summarized in Table 2. The OST were carried out as described by the original authors of the tests (Codman, 1934; Hawkins and Kennedy, 1980; Gerber et al., 1996; Liu et al., 1996; O'Brien et al., 1998; Kim et al., 2005; Barth et al., 2006).

During the belly-press test, peak isometric force was recorded and weakness was used as an additional criterion for a positive test result. Peak force was recorded using a hand-held dynamometer (Industrial Research Ltd, Christchurch, New Zealand) stabilized against the subjects' abdomen. The device was calibrated on the day of testing to within ±0.1 kg. Three trials were performed on each arm. The duration of the contraction was approximately 6–7 s and trials were followed by approximately 30 s rest.

Each examiner was blinded to the results of the other and there was no communication between examiners.

2.3. Statistical methods

Statistical analysis was carried out using Statistical Package for the Social Sciences (SPSS) version 16.0. Percent agreement between

Table 2
Criteria used for positive test result with orthopaedic special tests.

Test procedure	Test result	Criteria for a positive result
Active compression test:		
Acromioclavicular joint	+ve/−ve	Pain 'on top' of the shoulder (acromioclavicular joint) that was worse in the position of internal rotation, and relieved or abolished in the position of external rotation/supination.
Labral pathology	+ve/−ve	Pain or a click located 'inside' the shoulder that was worse in the position of internal rotation, and relieved or abolished in the position of external rotation/supination.
Hawkins–Kennedy test	+ve/−ve	Reproduction of subjects' symptoms
Drop-arm test	+ve/−ve	An inability to hold the arm at 90 degrees abduction, or a sudden drop of the arm when downward pressure is applied.
Crank test	+ve/−ve	Click produced during the test
Kim test	+ve/−ve	Production of posterior shoulder pain during the test
Belly-press test:		
Observation	+ve/−ve	Patient used shoulder extension to try to exert pressure resulting in elbow dropping behind body.
Weakness	+ve/−ve	Weakness of 30% or more compared with the opposite shoulder measured with a hand-held dynamometer (Industrial Research Ltd).

examiners, and Cohen's chance-corrected kappa statistics with associated 95% CI were calculated for the results of OST. Prevalence and bias adjusted kappa statistics (PABAK) were also calculated to account for unbalanced agreement category scores (prevalence) and differences in proportions of positive and negative results (bias) that are known to adversely affect overall kappa statistics (Landis and Koch, 1977; Feinstein and Cicchetti, 1990; Rigby, 2000; Shankar and Bangdiwala, 2008).

To determine reliability between examiners for the presence of weakness during the belly-press test, peak isometric force data was used to calculate a percent strength deficit of the affected side compared with the unaffected side, then converted to dichotomous values using a 30% or greater deficit as the criteria for a 'positive' response. Data from the peak force trial, and mean of three trials was used in the analysis. Two by two contingency tables were constructed for the results of the two examiners, and kappa statistics with associated 95% CI, PABAK and percent agreement statistics were calculated.

To determine whether extremes of prevalence or bias were likely to affect the overall kappa value, the prevalence index[1] (PI) and bias index[2] (BI) were calculated for each variable according to Byrt et al. (1993). PI values can range from −1 to +1, and the PI is equal to zero when 'yes' and 'no' are equally probable (Byrt et al., 1993). BI values can range from zero to 1, and equal to zero only if there is no difference in 'positive' proportions between examiners (Byrt et al., 1993).

Where the PI was high, we used the PABAK value for interpretation of results. For the purposes of this study, we determined an arbitrary cut-off value of a PI less than −0.5, or greater than 0.5, for interpretation of the PABAK values instead of overall kappa scores. Kappa and PABAK values for interexaminer reliability were interpreted according to the guidelines of Landis and Koch (1977); <0.00 poor; 0.00–0.20 slight; 0.21–0.40 fair; 0.41–0.60 moderate; 0.61–0.80 substantial; 0.81–1.00 almost perfect (Landis and Koch, 1977).

3. Results

Forty subjects with shoulder pain were recruited. Subjects included 23 males and 17 females with a mean age of 49 years (range 18–77 years). Descriptive data for subjects are summarized in Table 3. Randomization of examiner order resulted in 21 subjects being examined first by examiner 1 and 19 by examiner 2. Nine subjects with bilateral shoulder pain were excluded from the analysis of 'weakness' during the belly-press test.

Overall kappa values, raw prevalence of positive test results, PI, PABAK and percent agreement statistics for results of OST are presented in Table 4. BI results (not presented) ranged from 0.03 to 0.23 indicating the PABAK values were not significantly affected by examiner bias. On this basis, the PABAK values are interpreted as predominantly reflecting prevalence adjustment.

Interexaminer agreement on the results of OST was varied. Overall kappa values ranged from −0.04 to 0.65 ('poor' to 'substantial'). The prevalence of positive test results ranged from 15% (active compression test − acromioclavicular joint and Kim test) to 75% (active compression test − labral pathology). The PI exceeded −0.50 and 0.50 for a number of tests (active compression test − acromio-clavicular joint, drop-arm test, Kim test and belly-press test − observation and weakness) The PABAK values ranged from 0.65 to

---

[1] Prevalence Index = total number of 'positive' − total number of 'negative' agreements/number of cases.

[2] Bias Index = Difference in proportions of 'yes' for the two examiners/number of cases.

Table 3
Summary of subject characteristics.

Subject characteristics		Number	%
Gender	Male	23	58
	Female	17	42
Affected side	Dominant	24	60
	Non-dominant	11	28
	Bilateral	5	13
		Mean (Median)	Range
Age (years)		49 (51)	18–77
Height (cm)		171 (169)	157–189
Weight (kg)		80 (84)	53–102
Duration of symptoms (months)		48 (8)	<1–325
Pain severity in previous 24 h (11 point VAS)		3.6 (4.0)	0–7

0.78 for indicating 'substantial' agreement between examiners for these tests. Percent agreement ranged from 83% to 89%.

Tests where prevalence was not considered to adversely affect the kappa value included the active compression test (labral pathology), Hawkins–Kennedy test and crank test. 'Fair' agreement was demonstrated for these tests (kappa 0.36–0.38), and CI for the overall kappa values were wide. Percent agreement values for these tests ranged from 68 to 70%.

Highest levels of interexaminer agreement were observed for the belly press (weakness) (PABAK 0.78) and lowest agreement was observed for the crank test (kappa 0.36).

4. Discussion

We consider kappa or PABAK values in excess of 0.60, and percent agreement in excess of 80% are required for a test to be considered appropriate for inclusion in a clinical examination. The active compression test (acromioclavicular joint), drop-arm test, Kim test and belly-press test (observation and weakness), all reached this level in the present study.

Interexaminer reliability results for the drop-arm test, crank test in the current study are similar to those obtained where examiners were trained in orthopaedics and had a special interest in shoulders (Norregaard et al., 2002; Ostor et al., 2004; Nanda et al., 2008; Walsworth et al., 2008). In the present study, the prevalence of positive results for the drop-arm test was low, and further studies using larger numbers are required to confirm this result in the primary care environment.

Interexaminer reliability of the Hawkins–Kennedy test has been previously reported as 'fair' between an orthopaedic surgeon and a physical therapist (kappa 0.29; 95% CI 0.18, 0.40) (Razmjou et al., 2004). The only previous study identified involving symptomatic primary care patients in which both examiners were physiotherapists reported considerably higher reliability between examiners than observed in the current study (kappa 0.91; CI not reported) (Johansson and Ivarson, 2009). These authors investigated only four tests to identify subacromial pain. Differences in our results may be partially explained by the higher number of tests conducted in the present study, and the resulting potential for random error as a result of a change in the subjects' symptoms between assessments.

The results of the Kim et al. (2005) test in the current study (PABAK 0.70) differ from those of the original authors who reported an ICC value of 0.91. However, this study was conducted by the physician who developed the test and the methods and procedures for collection of interexaminer reliability data were not described. The Kim test demonstrated 'substantial' interexaminer reliability according to prevalence adjusted statistics, and high levels of

Table 4
Number of positive results, prevalence index, kappa, PABAK and percent agreement results.

Orthopaedic special test	Number of cases with positive result (% of total cases)	Prevalence index[a]	Overall kappa (95% CI)	PABAK[b]	% agreement
Active compression test					
Acromioclavicular joint	6 (15%)	0.83	0.22 (−0.24, 0.68)	0.75	88
Labral pathology	30 (75%)	−0.25	0.38 (0.1, 0.65)	0.40	70
Hawkins−Kennedy test	27 (68%)	−0.03	0.38 (0.10, 0.63)	0.35	68
Drop-arm test	7 (18%)	−0.78	0.57 (−0.14, 0.57)	0.67	83
Crank test	18 (45%)	−0.35	0.36 (−0.07, 0.59)	0.35	68
Kim test	6 (15%)	−0.85	−0.04 (−0.12, 0.03)	0.70	85
Belly-press test					
Observation	9 (23%)	−0.73	0.31 (−0.03, 0.64)	0.65	83
Weakness (maximal trial)	10 (25%)	−0.61	0.65 (0.33, 0.96)	0.78	89
Weakness (mean of 3 trials)	11 (28%)	−0.58	0.58 (0.26, 0.90)	0.72	86

[a] Prevalence Index = total number of 'positive' − total number of 'negative agreements/number of cases $(a − d/N)$.
[b] PABAK = $(2\,n/N) − 0.5 = 2p_o − 1/1 − 0.5$

agreement (85%) in the present study. Verification in a larger sample of the primary care population is required.

The results of our study provide the first known interexaminer reliability data for the belly-press test. Both components of the belly-press test (observation and weakness) demonstrated clinically acceptable levels of interexaminer reliability according to prevalence adjusted statistics (PABAK 0.65–0.78), and amongst the highest levels of raw agreement (83–89%) suggesting this is a reliable method of assessing the integrity of subscapularis.

The active compression test is reported to differentiate between acromioclavicular joint and glenoid labrum pathology (O'Brien et al., 1998). Previous studies indicate 'fair' reliability (kappa 0.24; 95% CI −0.02, 50) for the combined result of the active compression test for both pathologies (Walsworth et al., 2008). No studies were identified that tested the interexaminer reliability of differentiated test results of the active compression test for the two separate pathologies. Our results indicated a higher level of raw agreement between examiners (88%) and prevalence adjusted reliability (PABAK 0.75) in determining a positive result for acromioclavicular joint pain compared with labral pathology (raw agreement 70%; kappa 0.38). This finding may be explained by the relative ease with which subjects identify the more definitive, superficial pain localized to the "top" of the shoulder (acromioclavicular joint pain) compared with non-specific pain "inside" the shoulder.

Reported high pain severity and longstanding complaints have previously been identified as determinants of disagreement in diagnostic classification studies (de Winter et al., 1999). In this study, the mean duration of symptoms was high (48 months), which is longer than the duration of symptoms typically reported by patients in the primary care setting. Therefore the results of this study may not represent interexaminer reliability of these tests in patients with shoulder pain of shorter duration.

Several tests used in this study are reported to be diagnostic for shoulder pathologies that have a lower prevalence in the primary care population compared with orthopaedic settings, such as glenoid labrum tears (crank test, Kim test, active compression test) and subscapularis tears (belly-press test). This was reflected in the low prevalence of positive results for several of these tests in the present study. However the active compression test (labral pathology) and crank test were positive in 75% and 45% of our subjects respectively. The false positive and false negative rate of these and other tests has been investigated diagnostic validity study to determine the diagnostic accuracy of these tests for their respective pathology in the primary care population. Results will be reported in due course.

In conclusion, despite evidence of diagnostic accuracy, the active compression test (labral pathology), Hawkins−Kennedy test and crank test are unlikely to be of clinical value when high levels of agreement between examiners is required. The active compression test for acromioclavicular joint pathology, drop-arm test and Kim test demonstrated clinically acceptable levels of interexaminer agreement among primary care patients with sub-acute and chronic shoulder conditions. The diagnostic validity of these tests has been investigated in a recently completed study.

Acknowledgements

References

Barth JRH, Burkhart SS, De Beer JF. The Bear-Hug test: a new and sensitive test for diagnosing a subscapularis tear. Arthroscopy 2006;22(10):1076–84.

Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. Journal of Clinical Epidemiology 1993;46(5):423–9.

Codman EA. Rupture of the supraspinatus tendon and other lesions in or about the subacromial bursa. In: Todd T, editor. The shoulder; 1934. p. 123–77. Boston, Ch 5.

Cyriax JH. Textbook of orthopaedic medicine. In: Diagnosis of soft tissue lesions. 8th ed., vol. 1. Minneapolis: Balliere Tindall; 1982. pp. 1–50.

de Winter AF, Jans MP, Scholten JP, Deville W, van Schaardenburg D, Bouter LM. Diagnostic classification of shoulder disorders: interobserver agreement and determinants of disagreement. Annals of the Rheumatic Diseases 1999;58:272–7.

Feinstein AR, Cicchetti DV. High agreement but low Kappa: I. the problems of two paradoxes. Journal of Clinical Epidemiology 1990;43(6):543–9.

Fritz JM, Wainner RS. Examining diagnostic tests: an evidence-based perspective. Physical Therapy 2001;81(9):1546–64.

Gerber C, Hersche O, Farron A. Isolated rupture of the subscapularis tendon: results of operative repair. Journal of Bone and Joint Surgery (American) 1996;78A(7):1015–23.

Green R, Shanley K, Taylor NF, Perrott M. The anatomical basis for clinical tests assessing musculoskeletal function of the shoulder. Physical Therapy Reviews 2008;13(1):17–24.

Hawkins R, Kennedy J. Impingement syndrome in athletes. American Journal of Sports Medicine 1980;8:151–8.

Hegedus EJ, Goode A, Campbell S, Morin A, Tamaddoni M, Moorman CT, et al. Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. British Journal of Sports Medicine 2008;42:80–92.

Johansson K, Ivarson S. Intra- and interexaminer reliability of four manual shoulder maneuvers used to identify subacromial pain. Manual Therapy 2009;14(2):231–9.

Kim SH, Park JC, Jeong WK, Shin SK. The Kim test: a novel test for posteroinferior labral lesion of the shoulder − a comparison to the Jerk test. American Journal of Sports Medicine 2005;33(8):1188–92.

Landis JR, Koch GC. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

Liu SH, Henry MH, Nuccion S, Shapiro MS, Dorey F. Diagnosis of glenoid labral tears – a comparison between magnetic resonance imaging and clinical examinations. American Journal of Sports Medicine 1996;24(2):149–54.

Mimori K, Muneta T, Nakagawa T, Shinomiya K. A new pain provocation test for superior labral tears of the shoulder. American Journal of Sports Medicine 1999;27(2):137–42.

Murrell GAC, Walton J. Diagnosis of rotator cuff tears. Lancet 2001;357 (9258):769–70.

Nanda R, Gupta S, Kanapathipillai P, Liow RY, Rangan A. An assessment of inter examiner reliability of clinical tests for subacromial impingement and rotator cuff integrity. European Journal of Surgery and Traumatology 2008;18:495–500.

Norregaard J, Krogsgaard MR, Lorenzen T, Jensen EM. Diagnosing patients with longstanding shoulder joint pain. Annals of the Rheumatic Diseases 2002;61:646–9.

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. American Journal of Sports Medicine 1998;26(5):610–3.

Ostor AJK, Richards CA, Prevost AT, Hazleman BL, Speed CA. Inter-rater reliability of clinical tests for rotator cuff lesions. Annals of the Rheumatic Disorders 2004;63:1288–92.

Park HB, Yokota A, Gill HS, el Rassi G, McFarland EG. Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndrome. Journal of Bone and Joint Surgery (American) 2005;87(7):1446–55.

Razmjou H, Holtby R, Myhr T. Pain provocative shoulder tests: reliability and validity of the impingement tests. Physiotherapy Canada 2004;56(4):229–36.

Rigby AS. Statistical methods in epidemiology. V. Towards an understanding of the kappa coefficient. Disability and Rehabilitation 2000;22(8):339–44.

Shankar V, Bangdiwala SI. Behaviour of agreement measures in the presence of zero cells and biased marginal distributions. Journal of Applied Statistics 2008;35 (4):445–64.

Tokish JM, Decker MJ, Ellis HB, Torry MR, Hawkins RJ. The belly-press test for the physical examination of the subscapularis muscle: electromyographic validation and comparison to the lift-off test. Journal of Shoulder and Elbow Surgery 2003;12(5):427–30.

Walsworth MK, Doukas WC, Murphy KP, Mielcarek BJ, Michener LA. Reliability and diagnostic accuracy of history and physical examination for diagnosing glenoid labral tears. American Journal of Sports Medicine 2008;36(1):162–8.

Appendix 6

Manuscript Copy:

A Prospective Study of Shoulder Pain in Primary Care; Prevalence of Imaged Pathology and Response to Guided Diagnostic Blocks

Cadogan, A., Laslett, M., Hing, W., McNair, P., & Coates, M. (2011). A prospective study of shoulder pain in primary care: Prevalence of imaged pathology and response to guided diagnostic blocks. *BMC Musculoskeletal Disorders, 12*, 119. doi:10.1186/1471-2474-12-119

A prospective study of shoulder pain in primary care: Prevalence of imaged pathology and response to guided diagnostic blocks

Cadogan *et al.*

BMC
Musculoskeletal Disorders
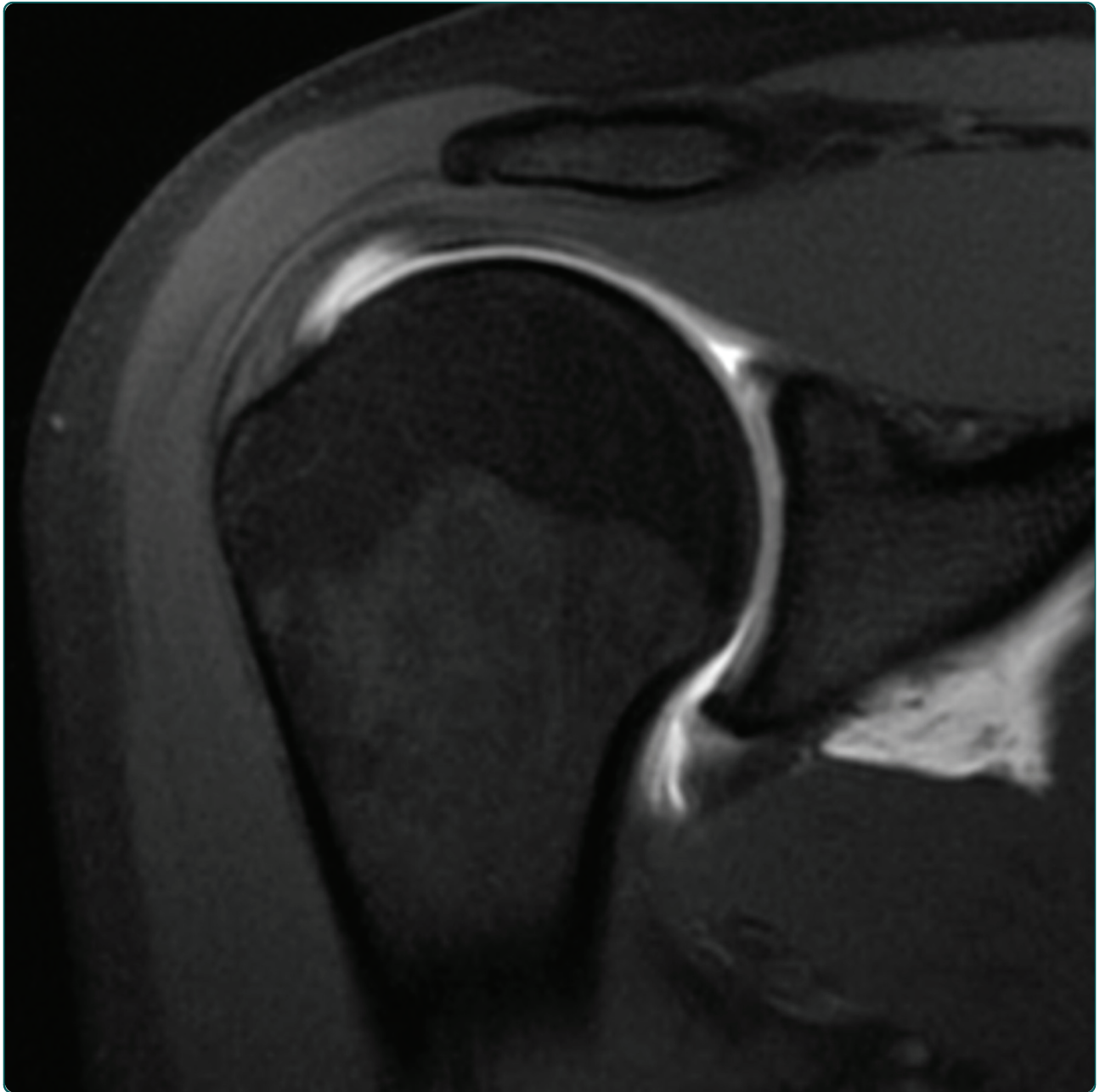
RESEARCH ARTICLE

Open Access

A prospective study of shoulder pain in primary care: Prevalence of imaged pathology and response to guided diagnostic blocks

Angela Cadogan[1*], Mark Laslett[1,2], Wayne A Hing[1], Peter J McNair[1] and Mark H Coates[3]

Abstract

Background: The prevalence of imaged pathology in primary care has received little attention and the relevance of identified pathology to symptoms remains unclear. This paper reports the prevalence of imaged pathology and the association between pathology and response to diagnostic blocks into the subacromial bursa (SAB), acromioclavicular joint (ACJ) and glenohumeral joint (GHJ).

Methods: Consecutive patients with shoulder pain recruited from primary care underwent standardised x-ray, diagnostic ultrasound scan and diagnostic injections of local anaesthetic into the SAB and ACJ. Subjects who reported less than 80% reduction in pain following either of these injections were referred for a magnetic resonance arthrogram (MRA) and GHJ diagnostic block. Differences in proportions of positive and negative imaging findings in the anaesthetic response groups were assessed using Fishers test and odds ratios were calculated a for positive anaesthetic response (PAR) to diagnostic blocks.

Results: In the 208 subjects recruited, the rotator cuff and SAB displayed the highest prevalence of pathology on both ultrasound (50% and 31% respectively) and MRA (65% and 76% respectively). The prevalence of PAR following SAB injection was 34% and ACJ injection 14%. Of the 59% reporting a negative anaesthetic response (NAR) for both of these injections, 16% demonstrated a PAR to GHJ injection. A full thickness tear of supraspinatus on ultrasound was associated with PAR to SAB injection (OR 5.02; $p < 0.05$). Ultrasound evidence of a biceps tendon sheath effusion (OR 8.0; $p < 0.01$) and an intact rotator cuff (OR 1.3; $p < 0.05$) were associated with PAR to GHJ injection. No imaging findings were strongly associated with PAR to ACJ injection ($p \leq 0.05$).

Conclusions: Rotator cuff and SAB pathology were the most common findings on ultrasound and MRA. Evidence of a full thickness supraspinatus tear was associated with symptoms arising from the subacromial region, and a biceps tendon sheath effusion and an intact rotator cuff were associated with an intra-articular GHJ pain source. When combined with clinical information, these results may help guide diagnostic decision making in primary care.

Background

Shoulder pain is a common and disabling complaint. The reported annual incidence of shoulder pain in primary care is 14.7 per 1000 patients per year [1] with a lifetime prevalence of up to 70% [2]. Recovery from shoulder pain can be slow and recurrence rates are high with 25% of those affected by shoulder pain reporting previous episodes, and 40 to 50% reporting persisting pain or recurrence at 12-month follow-up [3-5].

The most common causes of shoulder pain in primary care are reported to be rotator cuff disorders, acromio-clavicular joint (ACJ) disease and glenohumeral joint (GHJ) disorders [6], with classification of these disorders based primarily upon results of clinical tests [1,7-11]. However, inconsistent diagnostic terminology [12], lack of universally accepted diagnostic classification criteria [13,14] and poor specificity of many physical examination tests [15] hamper confidence in classification systems that use clinical test criteria alone.

Diagnostic imaging investigations including shoulder x-ray and diagnostic ultrasound imaging are increasingly being utilised by primary care practitioners to aid

* Correspondence: acadogan@vodafone.co.nz
[1]Health & Rehabilitation Research Institute, AUT University, Akoranga Drive, Private Bag 92006, Northcote, Auckland, New Zealand
Full list of author information is available at the end of the article

diagnosis [16]. More advanced imaging investigations such as magnetic resonance arthrogram (MRA) are also available, providing improved visualisation of pathologies such as glenoid labral lesions and tendon pathology [17]. While previous studies report the prevalence of imaging findings in the general population [18], specific athletic populations [19,20], samples of convenience [21,22] or case-control comparisons for specific shoulder pathology [23], the prevalence of imaged pathology in a prospective cohort of primary care patients suffering a current episode of shoulder pain has not been previously reported. Diagnostic decisions rely upon knowledge of prevalence of a condition in specific populations in order to estimate the likelihood of a positive 'disease' status or outcome following specific tests or investigations [24]. Knowledge of prevalence of imaged pathology in primary care would provide prior probability for specific conditions, thus assisting diagnostic decision-making processes and assessment as to the value of expensive or invasive investigations or interventions.

The interpretation of imaging findings can be complicated by the presence of anatomic variants [25,26] and the high prevalence of asymptomatic pathology especially in ageing populations [18,21]. The prevalence of asymptomatic full-thickness rotator cuff tears more than doubles after the age of 50 years [18], and asymptomatic ACJ arthritis has been identified by magnetic resonance imaging (MRI) in 93% of individuals over the age of 30 years [21]. Despite widespread use of imaging investigations in primary care, the relationship between imaging findings and symptoms has received limited attention. Diagnostic injections of local anaesthetic provide a method for determining whether symptoms arise from a specific structure [27,28]. Following injection of local anaesthetic into an anatomical structure, any subsequent reduction in pain intensity can be measured to assess the likelihood of its involvement in the patient's symptoms [29-31].

The aims of this paper were to report the prevalence of imaged shoulder pathology, and to evaluate the association between imaged pathology and a positive response to diagnostic blocks in a consecutive sample of patients with shoulder pain recruited from a primary care setting.

Methods
Study design and setting
The results presented in this paper formed part of a wider prospective, blinded diagnostic accuracy study in which clinical examination and imaging variables (index tests) were compared with results of diagnostic injections of local anaesthetic (reference standard) into the SAB, ACJ and GHJ. Subjects were recruited consecutively from a community-based medical centre and nine physiotherapy practices across Christchurch, New Zealand.

Ethical approval
The New Zealand Ministry of Health Regional Ethics Committee (Upper South A) granted ethical approval in May 2008.

Subjects
Consecutive patients presenting to their primary care practitioner (general practitioner (GP) or physiotherapist) for the first time with a new episode of shoulder pain (Figure 1), who were over 18 years of age and able to follow verbal instructions were eligible for inclusion in the study. Exclusion criteria were known fractures or dislocations around the shoulder complex, referred pain from the cervical spine, sensory or motor deficit involving the upper limb, previous surgery to the shoulder or cervical spine or contraindications to imaging or injection procedures.

Procedures
Subjects underwent a clinical examination (Additional file_1) followed by a standard shoulder x-ray series, diagnostic ultrasound scan and imaging guided diagnostic injections into the SAB and ACJ. Subjects reporting less than 80% reduction in pain intensity from either of these two injections were reviewed by a sports medicine physician prior to receiving an injection of local anaesthetic into the GHJ, performed as part of a contrast-enhanced MRA procedure. Study procedures are summarised in Figure 2.

X-ray and diagnostic ultrasound scan
Subjects underwent a standardized series of shoulder radiographs (x-ray) consisting of anterior-posterior (AP) views in neutral, external and internal rotation, axial view and outlet view [32]. X-rays were reported by experienced musculoskeletal radiologists. A standardised report form was used and radiologists recorded specific abnormalities of the ACJ, acromion, GHJ and calcific deposits. Imaging diagnostic criteria are presented in Table 1.

Diagnostic ultrasound scans were performed by trained and experienced musculoskeletal sonographers and reported by fellowship trained musculoskeletal radiologists. Examinations were performed using a Philips IU22 machine with a 5-12 MHz linear array probe using a standardised scan procedure [33,34]. The scan procedure is described in Additional file_2.

The SAB was observed during dynamic abduction and 'bunching' under the acromion and the coracoacromial ligament (CAL) was recorded. Subacromial bursal dimensions were measured from the deep margin of
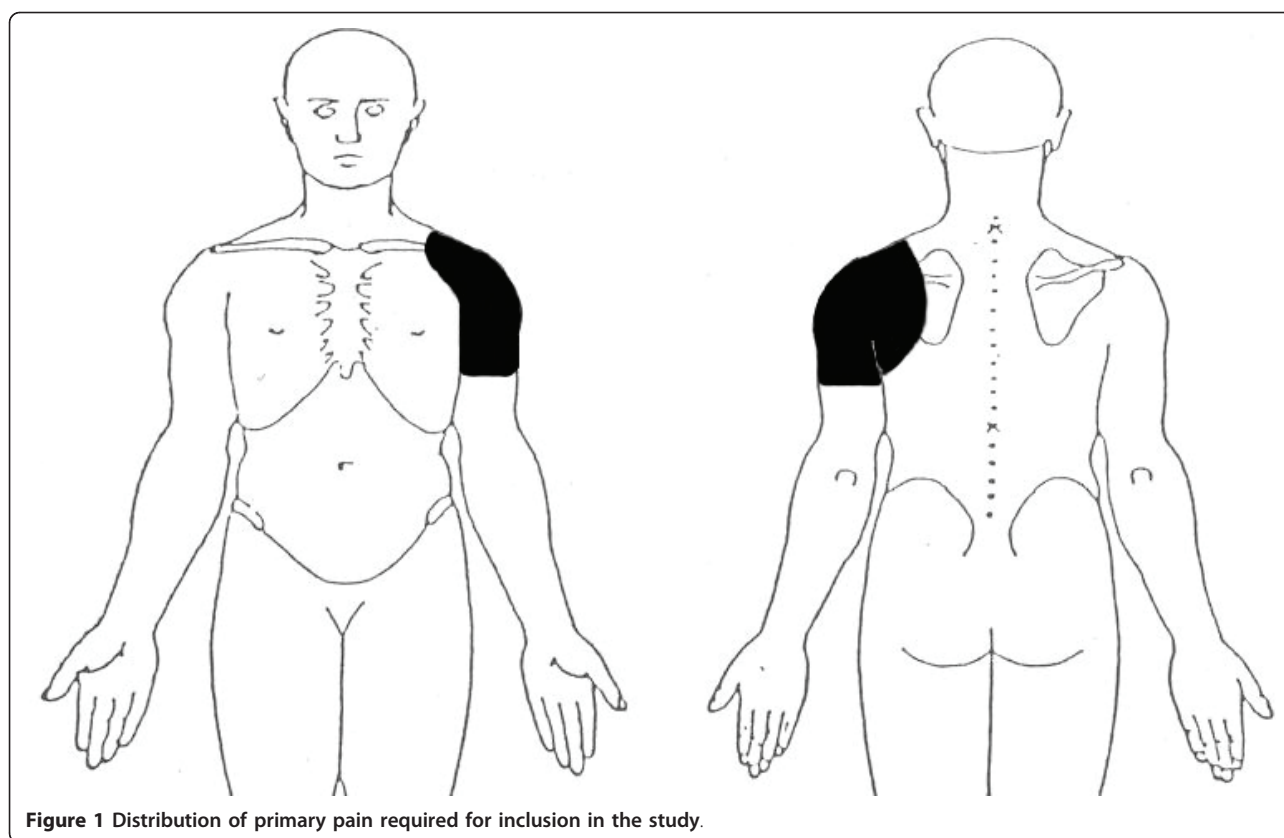
Figure 1 Distribution of primary pain required for inclusion in the study.

deltoid muscle to superficial margin of supraspinatus tendon in all cases where this distance was measurable (dimensions exceeding 1 mm).

Diagnostic injections
Subacromial bursa injection
Subjects were positioned supine with the arm in external rotation. Under aseptic conditions, a 22-gauge needle was used to inject 5 mL of 1% lidocaine hydrochloride (xylocaine™) into the SAB under ultrasound guidance using an anterior approach. When needle placement inside the SAB was confirmed by ultrasound, the contents of the syringe were emptied into the bursa. The radiologist recorded whether the SAB was successfully infiltrated. A video of this procedure may be viewed in Additional file 3_SAB injection, compatible with Windows® Media Player software.

Acromioclavicular joint injection
One week after the SAB injection, local anaesthetic was injected into the ACJ under fluoroscopic guidance using contrast enhancement. Subjects were positioned supine with the arm in external rotation. Under aseptic conditions, a 22-gauge needle was inserted into the ACJ using a direct anterior approach. Iodinated contrast (0.5 ml of Omnipaque 300 GE Healthcare) was introduced and fluoroscopic images used to confirm needle placement

within the ACJ. Approximately 2 mL of 1% lidocaine hydrochloride (xylocaine™) was then injected into the joint. The radiologist recorded whether the ACJ was successfully infiltrated and whether the injectate was contained within the joint. A video of this procedure may be viewed in Additional file 4_ACJ injection.

Glenohumeral joint injection
Approximately one week after the ACJ injection, subjects reporting less than 80% relief from both the SAB and ACJ injections underwent a GHJ arthrogram and intra-articular injection of local anaesthetic and gadolinium prior to magnetic resonance imaging (MRI). Subjects were positioned supine and the GHJ injection carried out under fluoroscopic guidance as described for the ACJ injection (above) using 5 mL of iodinated contrast. A mixture of 0.5 mL gadolinium (0.5 mmol/ml Gd-DOTA Guerbet France) and 10 mL 1% lidocaine hydrochloride (xylocaine™) was injected into the joint. The radiologist recorded whether the injectate was contained within the joint. A video of this procedure may be viewed in Additional file 5_GHJ injection.

Determination of post-injection change in pain intensity
Immediately prior to each injection, all subjects were examined using up to six clinical tests identified as being provocative of the subjects typical symptoms during the initial clinical examination (Additional file_1).
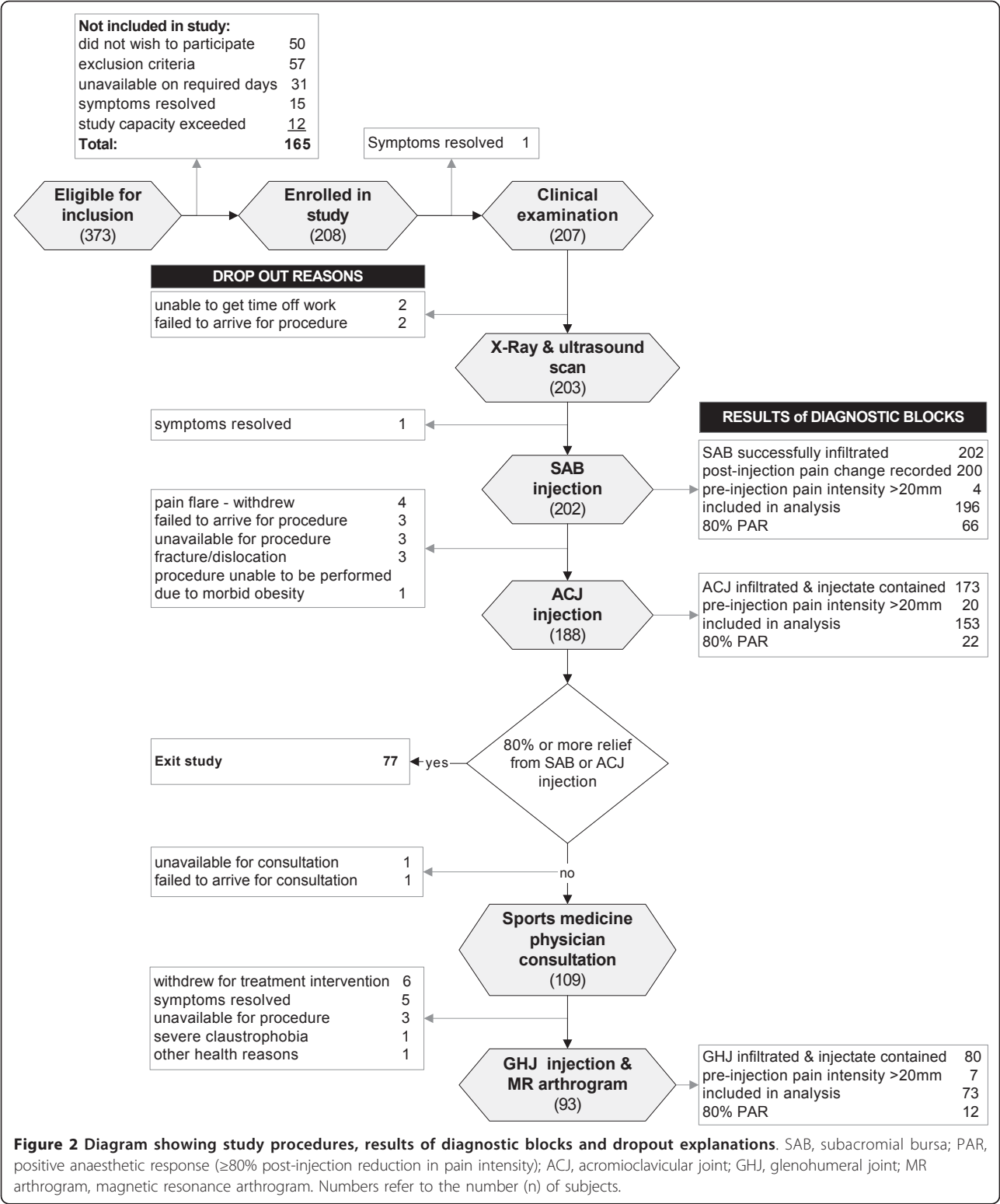
Figure 2 Diagram showing study procedures, results of diagnostic blocks and dropout explanations. SAB, subacromial bursa; PAR, positive anaesthetic response (≥80% post-injection reduction in pain intensity); ACJ, acromioclavicular joint; GHJ, glenohumeral joint; MR arthrogram, magnetic resonance arthrogram. Numbers refer to the number (n) of subjects.

Pre-injection pain intensity was recorded for each clinical test on a 100 m visual analogue scale (VAS) where 0 mm indicated "no pain" and 100 mm represented "worst imaginable pain". Tests were repeated between 5 and 15 minutes following each injection and post-injection pain intensity VAS scores recorded again. The percentage change in pain intensity (anaesthetic response) was calculated for each test [(post-injection VAS - pre-

Table 1 Imaging diagnostic criteria

Pathology	Imaging Diagnostic Criteria
X-Ray	
Acromioclavicular joint	
arthropathy/degenerative change	joint space narrowing, subchondral sclerosis, subchondral cystic change or marginal osteophytes.
osteolysis	bony resorption or increased lucency in distal clavicle.
Glenohumeral joint	
arthropathy/degenerative change	joint space narrowing, subchondral sclerosis, subchondral cystic change or marginal osteophytes.
other	loose bodies, joint calcifications.
Calcification of rotator cuff components	
supraspinatus	calcific deposits adjacent to the greater tuberosity on AP-external rotation x-ray view.
infraspinatus	calcific deposits adjacent to the greater tuberosity on AP-internal rotation x-ray view.
subscapularis	calcific deposits in the anterior shoulder region on axial x-ray view.
Ultrasound[a]	
ACJ pathology	Capsular hypertrophy, cortical irregularity or osteophytes, capsular bulge, joint space narrowing or widening.
Glenohumeral joint effusion	more than 2 mm between posterior glenoid labrum and posterior capsule.
Rotator cuff	
normal	normal contour, normal echogenicity.
calcification	focal increase in echogenicity with or without shadowing.
tendinosis	tendon thickening or decreased echogenicity.
tear	
intrasubstance	hypoechoic change not extending to articular or bursal surface.
partial thickness	SSp and ISp: hypoechoic change extending to either the articular or bursal surface. Subscapularis: partial fibre discontinuity.
full thickness	SSp and ISp: hypoechoic region extends from bursal to articular surface. Subscapularis: complete fibre discontinuity.
Subacromial bursa	
bursitis	hypoechoic fluid or effusion present and >1 mm thick.
bursal thickening	≥2 mm measured from deep margin of deltoid to superficial margin of supraspinatus.
"bunching"	Fluid distension of the SAB or 'buckling' of the rotator cuff during abduction
MR arthrogram[a]	
Acromioclavicular joint	
arthropathy/degenerative changes	capsular hypertrophy with or without joint space narrowing, subchondral cystic change, bone marrow oedema or osteophytes
osteolysis	bony resorption or bone marrow oedema in the distal clavicle
Rotator cuff	
normal	normal contour, normal signal
tendinosis	tendon thickening or mild increase in T2 signal
intrasubstance tear	linear increase in T2 signal which does not extend to the articular or bursal surface.
partial thickness tear	linear increase in T2 signal extending to the (bursal or articular) margins.
full thickness tear	fluid signal intensity or contrast extending from the bursal to the articular side lesion of the rotator cuff. Contrast seen in the SAB.
Subacromial bursitis	increased T2 signal within the SAB
Glenohumeral joint	
rotator interval pathology	thickening, signal change or tear involving the biceps pulley, superior glenohumeral or coracohumeral ligament, or synovitis in the rotator interval.
arthropathy/degenerative change	chondral loss, subchondral sclerosis, cystic changes, bone marrow oedema or osteophytes
labral tear	contrast extending into- or undermining the glenoid labrum, not conforming to normal variant anatomy.

Abbreviations: AP, antero-posterior view; ACJ, acromioclavicular joint; SSp, supraspinatus; ISp, infraspinatus; SAB, subacromial bursa;
[a]definitions based upon accepted diagnostic criteria [33,35]

injection VAS/pre-injection VAS)*100]. The average percent change from all tests was then calculated. A post-injection reduction in pain intensity of 80% or more was used as the criterion for a positive anaesthetic response (PAR). Subjects who did not reach an average of 80% pain relief following the SAB and ACJ injection were evaluated by a sports medicine physician and referred for the MRA investigation.

Magnetic resonance arthrogram imaging
Magnetic resonance imaging was obtained within 30 minutes of the GHJ injection. Imaging was performed with 3.0 Tesla General Electric-Milwaukee (GE) Signa HDxt platform running version 15 software. A conventional MR arthrography protocol was followed (Additional file_2) [35].

Blinding
The investigator performing the clinical examination and pre- and post-injection clinical tests (AC) was blinded to all diagnostic and treatment information from referring practitioners and to results of imaging procedures. Sonographers and radiologists were blinded to all clinical information prior to the x-ray, ultrasound scans and MRA procedure, and were blinded to results of anaesthetic response to injections.

Sample size considerations
Sample size was estimated using methods described by Flahault et al., (2005) [36]. Sample size was calculated for the diagnostic sub-group with the lowest expected prevalence (ACJ). The minimal acceptable lower confidence limit was set at 0.75 and expected sensitivity/specificity were both set at 0.90. A review of sample size estimates after the first 100 cases indicated lower than expected prevalence of PAR to ACJ diagnostic block and sample size was adjusted in order to maintain precision of diagnostic estimates.

Statistical analysis
The prevalence of imaged pathology and response to each of the diagnostic blocks are reported as frequency and percentages. Contingency tables (2×2) were constructed and Fishers exact test was used to compare proportions of positive and negative imaging findings in the anaesthetic response groups for each diagnostic injection procedure. *P*-values of ≤ 0.05 were used to indicate statistical significance. Odds ratios (OR) and 95% confidence intervals (CI) for PAR to diagnostic blocks were calculated. Statistical Package for the Social Sciences (SPSS) version 17.0 (IBM® Corporation 2010) was used for the analysis.

Due to the known limitations of VAS scales for measuring change in pain intensity when pre-injection pain levels are low (<20 mm) [37], only cases where pre-

injection pain intensity exceeded 20 mm were included in the analysis of anaesthetic response to diagnostic injections. Average percent change in pain intensity was calculated for the index tests with positive integers indicating increased post-injection pain intensity, and negative integers indicating decreased post-injection pain intensity.

Results
Subjects
A total of 208 subjects were included in the study between July 2009 and June 2010. Details of progression of subjects through the study and dropout explanations are presented in Figure 2. Demographic information for those included in the study is presented in Table 2. There were no significant differences between those included and excluded from the study with respect to age or gender. Symptom duration was shorter (median 2 weeks; IQ range 4 weeks) in subjects excluded from the study (Mann-Whitney *p* < 0.001). There were no significant differences in demographic characteristics between the total sample and the sub-group who received the GHJ injection as part of the MRA procedure (*p* > 0.05).

Prevalence of imaged pathology
X-ray and ultrasound scan
The prevalence of the pathologies identified on x-ray and ultrasound are presented in Figures 3 and 4. Acromioclavicular joint (Figure 5a) and GHJ pathology were the most common x-ray findings (both 17%) and calcification involving the rotator cuff was reported in 13% of subjects (Figure 5b).

Rotator cuff pathology was the most prevalent pathology on ultrasound (50%), with supraspinatus the most commonly affected rotator cuff component, accounting for 86 of the 102 cases (85%) of rotator cuff pathology. Tears were the most common pathology affecting supraspinatus accounting for 52% of all supraspinatus pathology and intrasubstance tears were the most common type of tear accounting for 51% of all supraspinatus tears (Figure 6a). Calcification was the most common finding in infraspinatus (59%) and subscapularis (69%) compared with 39% in supraspinatus.

Prevalence of SAB pathology was 31% and bursal thickening (dimensions exceeding 2 mm) was reported in 23% of subjects (Figure 6b). Bunching of the SAB under the acromion was observed in 84 subjects (43%) (Figure 6c), and this was associated with reproduction of symptoms in 72 subjects (86% of cases in which bunching was observed). Bunching under the CAL was observed in 51 of the 94 cases (54%) in which this was assessed, and was associated with reproduction of symptoms in 40 subjects (78% of cases in which bunching was observed) (Figure 7).

Table 2 Subject demographics

Subject characteristics	All subjects (n = 208)		MRA group (n = 93)	
	Mean (SD)	Range	Mean (SD)	Range
Age (years)	42 (14)	18-81	42 (14)	18-81
Height (cm)	172 (10)	147-199	172 (10)	151-198
Weight (kg)	80.6 (18.0)	50.3-189.0	82.3 (15.8)	52.7-125.3
Symptom duration (weeks)*	7 (13)	0-175	7 (13)	0-175
Worst pain previous 48 hours (100 mm VAS)	62 (23)	3-100	63 (24)	3-100
Average pain previous 48 hours (100 mm VAS)	37 (22)	1-100	37 (24)	1-100
	n (%)		**n (%)**	
Male gender	107 (51)		53 (57)	
Right hand dominant	110 (53)		79 (85)	
Dominant arm affected	110 (53)		48 (52)	
ACC Claim	193 (93)		86 (93)	
Referrals				
physiotherapist	203 (98)		89 (96)	
general practitioner	5 (2)		4 (4)	
Employment status				
in paid employment	166 (80)		76 (82)	
on modified duties due to shoulder pain	18 (9)		10 (11)	
off work due to shoulder pain	7 (3)		4 (4)	
not currently employed/working	41 (20)		17 (18)	
Co-existent medical conditions	70 (34)		33 (36)	
Current smoker	39 (19)		18 (20)	

Abbreviations: MRA, magnetic resonance arthrogram; SD, standard deviation; VAS, visual analogue scale: ACC, Accident Compensation Corporation.
*symptom duration was not normally distributed. Figures presented are median (IQ range).

Magnetic resonance arthrogram

The prevalence of MRA findings is shown in Figure 8. Only one case was reported as "normal" (no abnormality reported) and 74% of cases demonstrated multiple pathologies. The most commonly reported MRA finding overall was SAB pathology (76%) with subacromial bursitis reported in 68 subjects (73%) (Figure 9a). Rotator cuff pathology affected at least one of the rotator cuff components in 65% of cases. Supraspinatus was the most frequently affected component of the rotator cuff (85% of all rotator cuff pathology) and tears were the most common pathological finding in all rotator cuff components accounting for 41 of the 61 cases (67%) of rotator cuff pathology. Partial thickness tears involving the articular surface were the most common type of supraspinatus tear identified (34% of all supraspinatus tears) (Figure 9b). GHJ pathology (63%) and ACJ pathology (59%) were also highly prevalent with rotator interval pathology (GHJ) and degenerative ACJ changes (Figure 9c) (both 55%) the most common findings. Glenoid labrum tears were present in 47% of all subjects who received the MRA and were associated with paralabral cysts in 10 cases (23%). Suprascapular nerve compression was associated with paralabral cysts in two cases (2%) (Figure 9d).

Prevalence of anaesthetic response to diagnostic blocks

The anaesthetic response profiles for the diagnostic injections are presented in Figure 10. There were no observable differences in the frequency of imaged pathology between those in whom post-injection pain intensity increased compared with cases in which a post-injection decrease in pain was reported. Results for the injection procedures are presented in Figure 2. Infiltration of the SAB was confirmed in all cases and a PAR ($\geq$80% pain relief) was reported by 66 subjects (34%) following the SAB injection. Average ACJ injection volume was 2.1 mL (SD 0.7 mL) and 22 of the 153 subjects (14%) in whom the injectate was contained within the ACJ and whose pre-injection pain intensity exceeded 20 mm on the 100 mm VAS scale reported an 80% PAR. Ninety three subjects received the GHJ injection as part of the MR arthrogram procedure and an 80% PAR was reported by 12 of the 75 subjects (16%) in whom the injectate was contained within the GHJ and pre-injection pain intensity exceeded 20 mm.

Association between imaged pathology and response to diagnostic blocks

Imaging variables associated with PAR to diagnostic block ($p \leq 0.05$) and demonstrating a magnitude of

Figure 3 Prevalence of pathology identified on x-ray. n, number of cases; ACJ, acromioclavicular joint; GHJ, glenohumeral joint
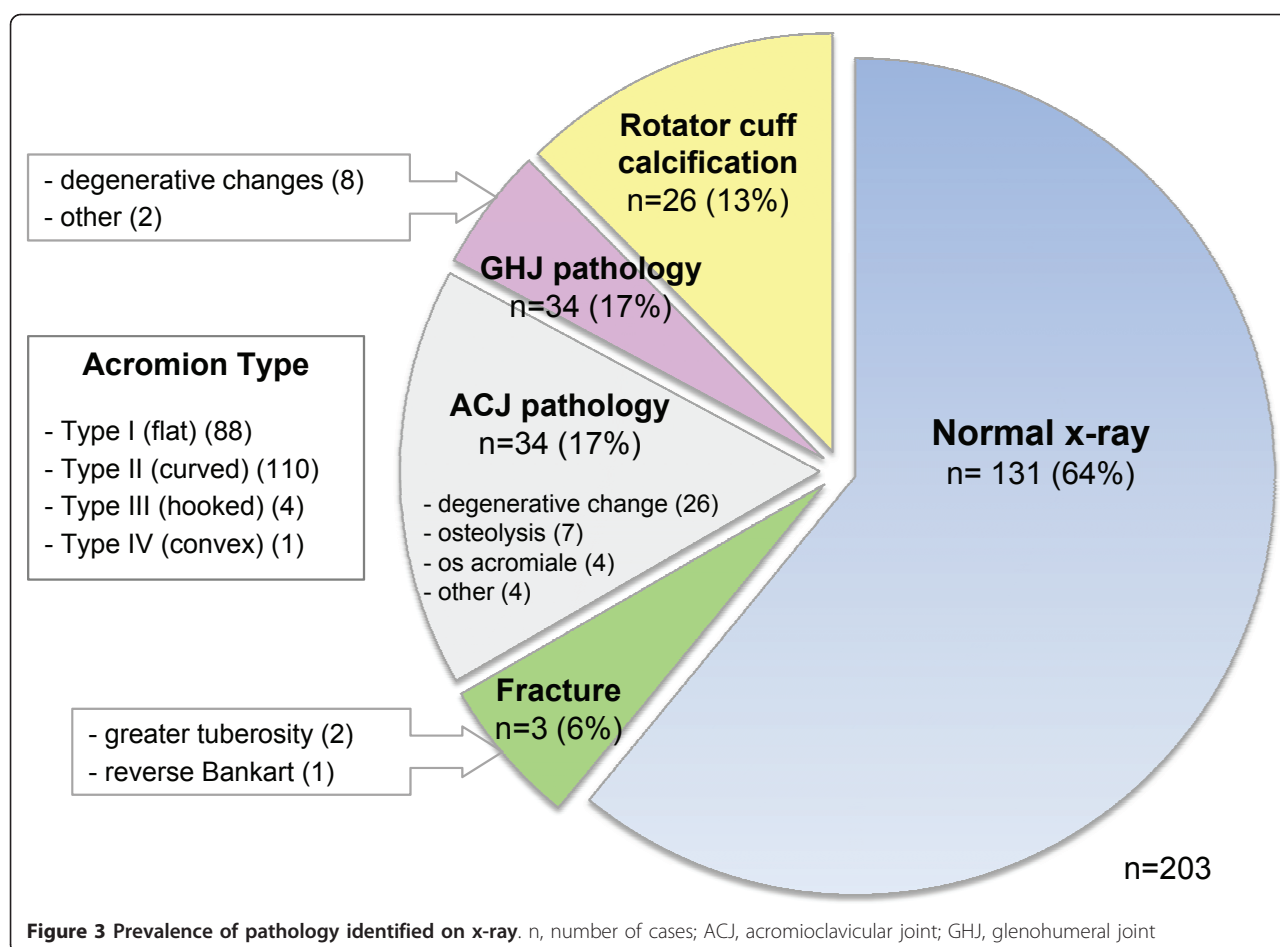
association OR greater than 2.0 are summarised in Table 3. Results for all other x-ray and ultrasound variables are presented in Additional file_6 (SAB and ACJ injection) and Additional file_7 (GHJ injection). Results for all other MRA variables are presented in Additional file_8.

A full thickness supraspinatus tear identified by ultrasound imaging was associated with PAR to SAB injection (OR 5.0, $p \leq 0.05$). None of the imaging variables were strongly associated with PAR to ACJ injection ($p > 0.05$). The strongest association of any imaging variable with diagnostic block was the association between biceps tendon sheath effusion identified on ultrasound and PAR to GHJ injection (OR 8.0; $p < 0.01$). A tear of the rotator cuff reported on ultrasound was negatively associated with a PAR to GHJ injection ($p < 0.05$). When recoded, an 'intact' rotator cuff on ultrasound demonstrated an OR of 1.3 for a PAR.

Discussion

This is the first report of the prevalence of imaged pathology and anaesthetic responses to diagnostic injection into the SAB, ACJ and GHJ in a sample of primary care

patients with shoulder pain. Estimates of the likelihood of symptomatic pathologies being present that affect these sites will increase or decrease as details from the history and physical examination are added to the imaging findings, but prior probability (prevalence) of these conditions in the population of interest is the necessary baseline and starting point [24]. This study provides the prior probability data for specific pathologies and pain sources at the 80% pain reduction level in a sample of primary care patients. This knowledge may help inform clinical decisions regarding treatment interventions, the use of advanced imaging or specialist referral.

Prevalence of imaged pathology
X-ray and diagnostic ultrasound scan

Shoulder x-rays were reported as 'normal' in 64% of cases however the detection of three unsuspected fractures in our study population highlights the use of x-ray as a valuable screening tool. The prevalence of calcification identified on x-ray (13%) was similar to previous reports (10%) [38].

Subacromial bursa pathology was a common ultrasound finding (31%) in our symptomatic sample. We
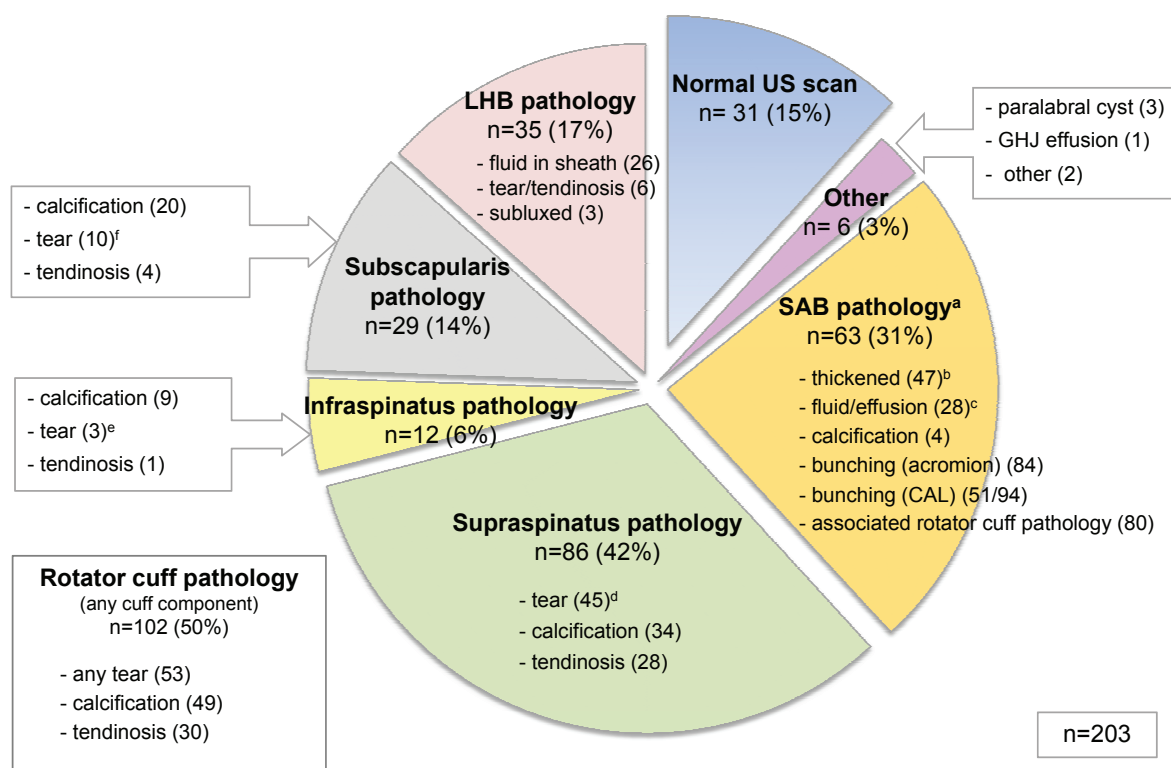
Figure 4 Prevalence of pathology identified on ultrasound scan. (n), number of cases; US, ultrasound; GHJ, glenohumeral joint; SAB, subacromial bursa; CAL, coracoacromial ligament; LHB, long head of biceps tendon. [a]Subacromial pathology: any one of three present; dimension ≥2 mm, fluid/effusion or calcification. [b]Subacromial bursa dimensions: <1 mm (71); 1-2 mm (82); 2-3 mm (42); >3 mm (5). [c]Subacromial bursal effusion associated with full thickness rotator cuff tear (7). [d]Supraspinatus tears: intrasubstance (23); partial thickness-bursal surface (4); partial thickness-articular surface (8); full thickness (10). [e]Infraspinatus tears: intrasubstance (1); partial thickness (1); full thickness (1). [f]Subscapularis tears: intrasubstance (5); partial thickness (4); full thickness (1).

Figure 5 Shoulder x-ray images of ACJ pathology and rotator cuff calcification. a) AP x-ray view in external rotation showing degenerative acromioclavicular joint changes (white arrow); b) outlet view showing calcification in line with the infraspinatus tendon (black arrow).

Figure 6 Ultrasound scan images of subacromial bursa and supraspinatus pathology. a) hypoechoic region (between calipers) indicating an intrasubstance tear within posterior fibres of supraspinatus (longitudinal view) overlying the head of humerus (white arrowhead); b) thickened subacromial bursa (calipers); c) bunching of the SAB (white arrow) under the acromion during dynamic abduction.

used the criterion of bursal dimension ≥2 mm, calcification or bursal fluid or effusion or to classify 'SAB pathology'. Opinions vary regarding the dimension (thickness) at which the normally thin hypoechoic line of the SAB is regarded as pathological. Some have suggested the ability to view and measure the SAB at all represents pathological thickening [39], others consider more than 2 mm thickness to be pathological [40-42] and some suggest SAB thickness compared with the unaffected side irrespective of bursal dimension to be of more clinical relevance [43]. Recent theories question whether SAB thickening is even pathological, proposing it may be the result of adaptation to repeated overhead

activity [16]. Variable agreement (kappa 0.50 to 0.89) has also been reported between musculoskeletal ultrasound experts for identification of SAB pathology on ultrasound [44-47] with most disagreements relating to variations in dynamic assessment and judgement of SAB fluid as being normal or pathological [47]. Technicalities surrounding the ultrasound diagnosis of SAB pathology, lack of expert consensus upon the dimension at which the SAB is considered pathological and the poor understanding of the relationship between SAB histopathology and imaging findings mean that the reported prevalence of SAB pathology on ultrasound is likely to vary. Bursal bunching was also identified in a high proportion of

	Acromion (n=195)	Coracoacromial ligament (n=94)
symptomatic bunching	37%	42%
asymptomatic bunching	6%	12%
no bunching	57%	46%
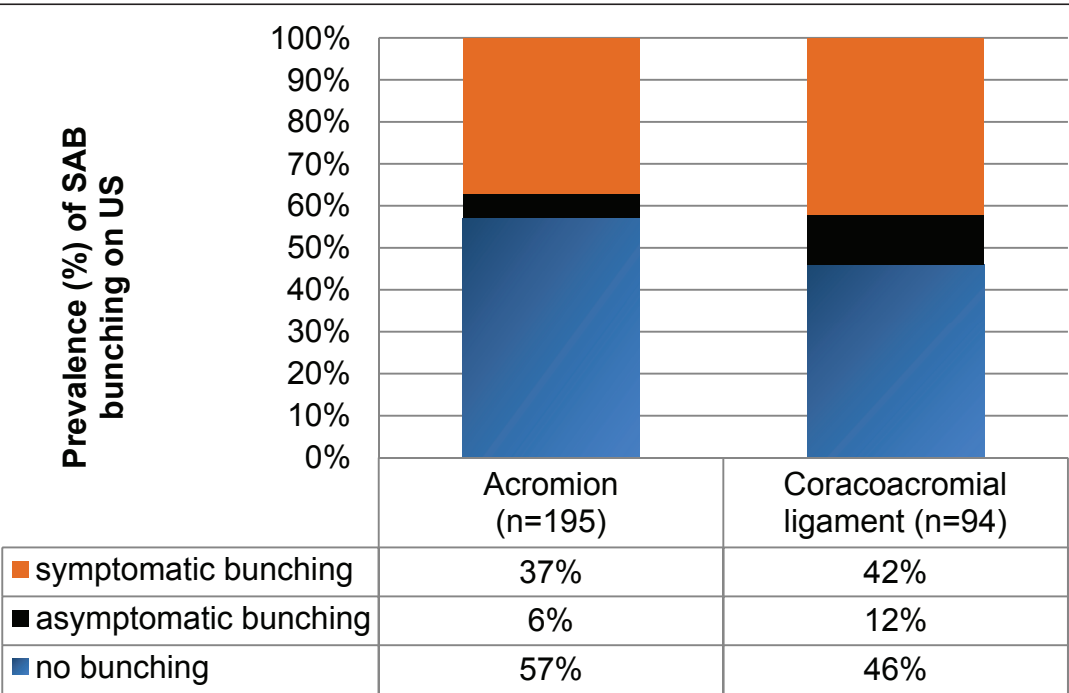
Figure 7 Prevalence of subacromial bursa bunching under the acromion and coracoacromial ligament on ultrasound during dynamic abduction. SAB, subacromial bursa; US, ultrasound; CAL, coracoacromial ligament. Percentages are in reference to the number of cases in which bursal bunching was assessed (acromion n = 195; CAL n = 94).

Figure 8 Prevalence of pathology identified on MR arthrogram. (n), number of cases; LHB, long head of biceps tendon; ACJ, acromioclavicular joint; GHJ, glenohumeral joint; OA, osteoarthritis; SAB, subacromial bursa; [a]ACJ degenerative changes: mild (28); moderate (18); severe (5). [b]Acromioclavicular joint pathology - other: os acromiale (2); unfused acromial ossification centre (1); acromial spur (4); widened joint space/subluxation (2); synovitis (1). [c]Rotator interval pathology: coracohumeral or superior glenohumeral ligament thickening (40); rotator interval synovitis (39); biceps pulley, coracohumeral or superior glenohumeral ligament tear (13). [d]Glenoid labrum tear: isolated labral tear (5); associated pathology present (39); SLAP tear (20); SLAP Type II (17), Type III (2), Type IV (1); anterior-inferior tear (9); semi- or full circumferential tear (7); posterior-superior tear (1); other tear (9); paralabral cyst (10); paralabral cyst causing suprascapular nerve compression (2). [e]Glenohumeral joint pathology - other: bony irregularity humeral head without marrow oedema (12); Hill-Sachs lesion (3); intra-articular/osseous body (3); ganglion cyst between coracoacromial and coracohumeral ligaments (1); greater tuberosity fracture (1). [f]Subacromial bursitis: mild (52); moderate (12); severe (4) [g]Supraspinatus tears: intrasubstance (11); partial thickness-bursal surface (5); partial thickness articular surface (12); full thickness (7). [h]Infraspinatus tears: intrasubstance (4); partial thickness (3); full thickness (0) [i]Subscapularis tears: intrasubstance (4); partial thickness (0); full thickness (2)

subjects, however bunching was asymptomatic in 14% (acromion) and 22% (CAL) of cases in which bunching was observed. This highlights the need to correlate imaging findings with clinical symptoms when considering the diagnosis of 'subacromial impingement'.

Magnetic resonance arthrogram
Magnetic resonance arthrogram findings in the subgroup of subjects receiving the investigation, revealed a high prevalence of multiple pathologies (74%), similar to previous reports (77%) in an asymptomatic primary care population [48]. In the subjects who received the MRA, SAB and ACJ pathology were reported respectively in 76% and 59% of subjects, all of whom had previously been classified as 'non-responders' at the 80% pain relief level following injection of local anaesthetic into these

structures. Marrow oedema on MRI has been reported as a reliable indicator of symptomatic ACJ pathology [23]. Our study identified eight cases (9%) of active ACJ arthropathy with marrow oedema in subjects who had previously demonstrated a NAR to ACJ injection, however the inability of the local anaesthetic to penetrate to the level of subchondral bone, thereby classifying those subjects as 'non-responders' to ACJ injection, represents a likely explanation for this result.

Rotator cuff pathology was reported in more than half of subjects on both ultrasound and MRA with rotator cuff tears identified in 26% and 44% of subjects with the respective imaging procedures. Although no primary care imaging studies are available for direct comparison, these results are similar to previous reports of the
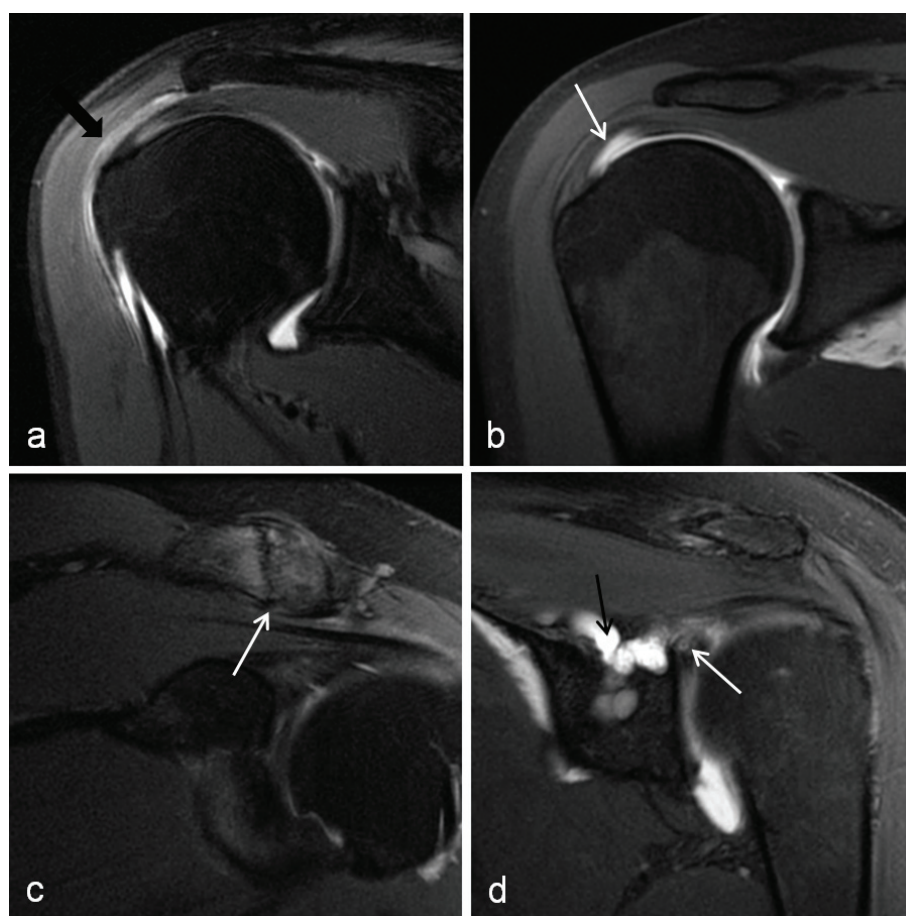
Figure 9 MR arthrogram images of shoulder pathology. a) subacromial bursitis - coronal PD fat saturated image showing region of hyperintensity in the subacromial bursa (black arrow); b) partial thickness, articular surface supraspinatus tear (white arrow) - coronal T1 fat saturated image showing contrast extending into the supraspinatus tendon. c) ACJ degenerative changes (white arrow) -coronal PD fat saturated image; d) type III SLAP tear (white arrow) with contrast filling a paralabral cyst (black arrow) which extended into the supraglenoid and suprascapular notch causing neural compression -coronal PD fat saturated image.

prevalence of rotator cuff tears in asymptomatic populations on ultrasound [18] and MRI [48]. Of interest was the higher number of intrasubstance tears involving infraspinatus, and partial thickness (articular surface) supraspinatus tears identified on MRA compared with the number identified on ultrasound imaging, despite the smaller sample number in this subgroup. While identification of an intrasubstance tear on MRA is unlikely to alter management at primary care level unless it is associated with more serious pathology, partial thickness tears of the rotator cuff are reported to be of prognostic significance due to the high proportion that increase in size or progress to full thickness tears if left untreated [49]. Ultrasound imaging has previously demonstrated only moderate pooled sensitivity (72%) for detection of partial thickness rotator cuff tears compared with MRI or surgery [50]. Variable agreement among experts on the presence of partial thickness

rotator cuff tears on ultrasound (kappa 0.63; 88% to 92% agreement) has also been reported [44,47,51]. Results of MRI scans have been shown to alter clinical decisions regarding management of rotator cuff tears in the orthopaedic setting [52] and MRA may therefore be indicated at the primary care level if there is clinical suspicion of rotator cuff disruption in the presence of equivocal ultrasound findings.

The prevalence of intra-articular GHJ pathology on MRA in this sub-group of subjects was also high (63%) with rotator interval pathology (55%) and glenoid labral tears (47%) the most common findings. However, despite the high prevalence of GHJ pathology in this study, only 16% of individuals were classified as responders to the GHJ injection at the 80% pain relief level. During the MRA procedure, contrast was introduced into the GHJ through the region of the rotator interval and in some subjects the appearance of contrast in this

Figure 10 Anaesthetic responses to diagnostic blocks. SAB, subacromial bursa; ACJ, acromioclavicular joint; GHJ, glenohumeral joint.
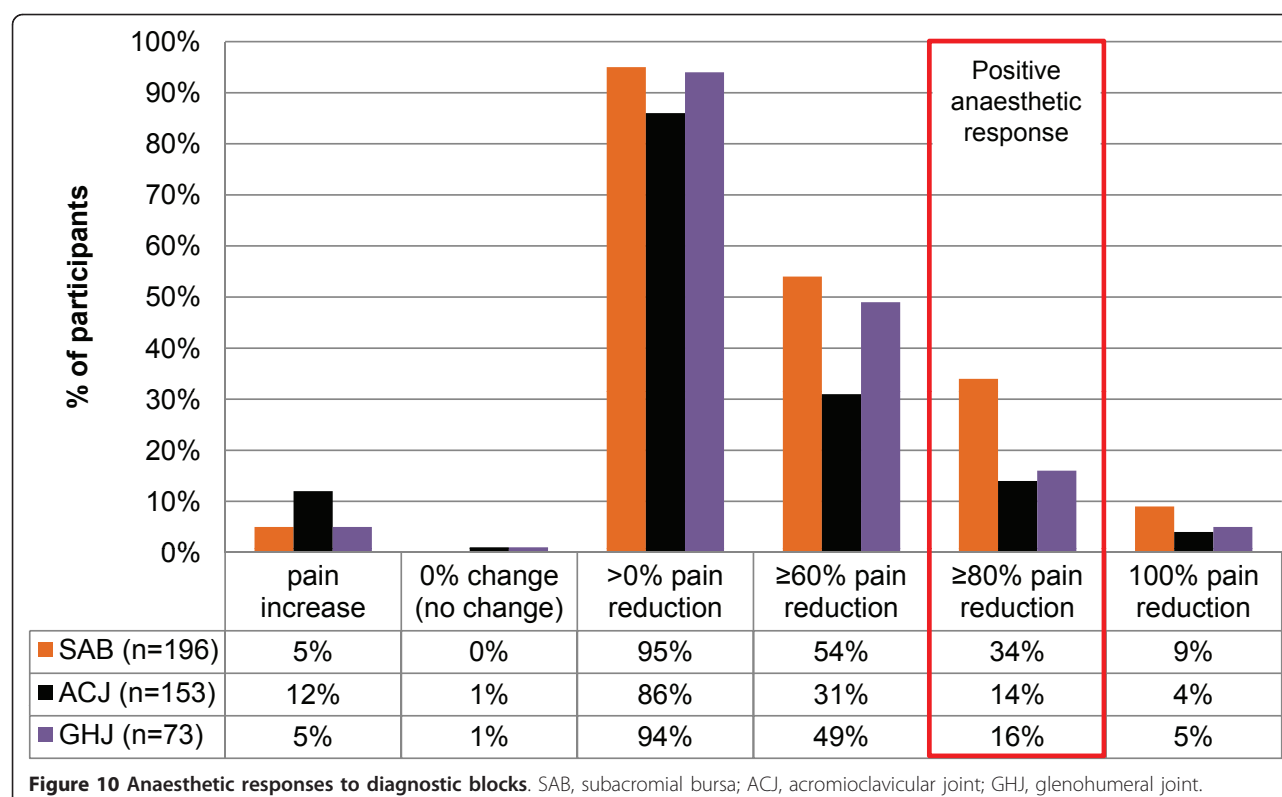
Table 3 Summary of imaging variables demonstrating association with positive anaesthetic response to diagnostic blocks (*p* ≤ 0.05 or OR >2.0)

Pathology identified on imaging	Pathology identified (total cases) (n)	% with pathology present reporting PAR	% with pathology absent reporting PAR	OR (95% CI)	Fishers test (*p* value)
SAB injection (PAR n = 66)					
X-ray: type 3 acromion	4	75	33	6.2 (0.64, 61.23)	0.109
X-ray: os acromiale	4	75	33	6.1 (0.63, 60.25)	0.112
X-ray: supraspinatus calcification	16	56	31	2.8 (1.00, 7.97)	0.054
US: supraspinatus calcification	33	49	31	2.1 (1.00, 4.55)	0.068
US: supraspinatus FTT	10	70**	32	5.0 (1.25, 20.11)	0.033
ACJ injection (PAR n = 22)					
X-ray: ACJ pathology	21	14	16	2.1 (0.69, 6.52)	0.189
US: supraspinatus tear PTT (articular surface)	8	0	17	2.1 (0.39, 11.05)	0.323
US: LHB tendinosis	3	0	16	3.1 (0.27, 35.39)	0.374
GHJ injection (PAR n = 12)					
US: no rotator cuff tear	19	21**	0	1.3 (1.11, 1.46)	0.029
US: supraspinatus tendinosis	11	27	14	2.3 (0.51, 10.30)	0.374
US: subscapularis tendinosis	3	33	15	2.8 (0.23, 33.27)	0.421
US: biceps tendon sheath effusion	13	46**	10	8.0 (2.02, 31.72)	0.004
MRA: ACJ pathology	46	20	11	2.0 (0.50, 8.23)	0.516
MRA: osteolysis lateral clavicle	5	40	15	3.9 (0.58, 26.58)	0.187
MRA: contrast seen in SAB	6	33	15	2.9 (0.47, 17.99)	0.254

Abbreviations: PAR, positive anaesthetic response (≥80% post-injection pain intensity reduction); OR, adjusted odds ratio; CI, confidence interval; SAB, subacromial bursa; US, ultrasound; FTT, full thickness tear; ACJ, acromioclavicular joint; PTT, partial thickness tear; LHB, long head of biceps; GHJ, glenohumeral joint; MRA, magnetic resonance arthrogram.

Percentages do not total 100% as these represent proportion of subjects with or without pathology on imaging (row percentages in contingency table) in the PAR group. Negative anaesthetic response group results (column percentages) are not presented.

**p ≤ 0.05

region on subsequent MRI films may have been difficult to distinguish from mild rotator interval pathology. Glenoid labral tears are frequently associated with other extra-articular pathology such as rotator cuff tears [53-56], and the rotator interval also has complex pathoanatomic relationships with supraspinatus, subscapularis and the long head of biceps tendon [57]. The high proportion of multiple pathology and low GHJ PAR rate in this study may be partially explained by the concurrent involvement of extra-articular structures.

Association between imaging findings and anaesthetic response

Subjects with full thickness tears of supraspinatus identified by ultrasound imaging were more likely to experience a PAR to SAB injection than those without a full thickness tear. Full thickness supraspinatus tears affect the SAB-rotator cuff interface and infiltration of the torn cuff with anaesthetic through this disruption is the likely explanation for this finding. The small proportion of PAR among those with an intrasubstance supraspinatus tear (intact margins) reported on ultrasound supports this theory, however none of the four cases in which bursal-surface supraspinatus tears were identified were classified as responders to the SAB injection. None of the imaging variables were strongly associated with PAR to ACJ injection. The high prevalence of asymptomatic degenerative changes particularly in individuals older than 30 years (93%) [21] may explain this result.

A long head of biceps tendon sheath effusion on ultrasound was significantly related to a PAR to GHJ injection. The biceps tendon sheath is a synovial extension of the GHJ capsule and may therefore be indicative of a GHJ effusion resulting from intra-articular GHJ pathology or systemic inflammatory disease. A biceps tendon sheath effusion on ultrasound has been shown to be more sensitive than arthrography for detection of intra-articular GHJ pathology [58]. It is also a common finding in those suffering rheumatoid arthritis [59,60] and has been found to be predictive of degenerative GHJ arthritis and polymyalgia rheumatica [53,61]. In the current primary care study, half the subjects with a biceps tendon sheath effusion reported on ultrasound were classified as positive 'responders' to the GHJ diagnostic block at the 80% pain reduction standard. The likely explanation for the PAR is the anaesthetisation of synovial tissue within the GHJ. Although this finding may implicate an intra-articular pain source, it is a non-specific result and further imaging investigations such as MRI or laboratory tests would be required to identify the specific pathology responsible for the synovial effusion. The magnitude of association of the biceps tendon sheath effusion on ultrasound with PAR to GHJ injection seen in this study (OR 8.00), and a lower 95% confidence limit of 2.0 suggest this finding may be of value in the primary care setting when considering further imaging investigation, laboratory testing or referral for higher levels of care.

Subjects with an intact rotator cuff on ultrasound also demonstrated a higher proportion of PAR to GHJ injection ($p < 0.05$) than those in whom a rotator cuff tear was identified. This could imply that in subjects with a rotator cuff tear, the tear itself may have been more symptomatic than any co-existent intra-articular GHJ pathology resulting in the NAR to GHJ diagnostic block. Although the OR for PAR to GHJ injection in the presence of an intact rotator cuff on ultrasound was small (1.27), the CI did not include 1.0, and could represent a clinically meaningful increase in the likelihood of a PAR since the prevalence of this imaging finding was high (74%) [62]. Current guidelines advocate ultrasound imaging only when a major rotator cuff tear is suspected when surgery may be considered as a treatment option [63]. However, these results may provide additional justification for the use of diagnostic ultrasound imaging in the primary care setting to inform decisions regarding further investigations for intra-articular GHJ pathology in the presence of an intact rotator cuff and relevant clinical findings.

Limitations of the study

The definition of 'accident' in the context of subject 'claim status' in this study is influenced by New Zealand's' unique Accident Compensation Corporation legislation. Although the majority of subjects included in our study had a current ACC claim, this does not necessarily imply a significant degree of trauma, and complaints included many less severe conditions with low levels of functional disability. Those whose shoulder pain is not covered by an ACC claim may, however, be less likely to present for medical assessment and may be under-represented in this study. Due to the cost of the MRA procedures it was not possible for every subject to undergo this procedure, and several subjects with high and low levels of pain intensity withdrew from the study prior to the MRA representing a potential source of selection bias in this subgroup of subjects.

Conclusions

Rotator cuff and SAB pathology were the most common findings on both ultrasound and on MRA in this primary care cohort. A full thickness supraspinatus tear on ultrasound was associated with subacromial pain according to our criterion, and ultrasound findings of a biceps tendon sheath effusion and an intact rotator cuff were associated with pain arising from the GHJ in a subgroup of subjects. Results provide the prior probability of imaged pathology, and when combined with clinical examination findings may inform decisions in primary care regarding treatment interventions and the need for advanced diagnostic imaging or specialist referral.

Consent statement

Written informed consent was obtained from the patient for publication of this case report and accompanying images. A copy of the written consent is available for review by the Editor-in-Chief of this journal.

Additional material

Additional file 1: Clinical examination procedures. Table listing the clinical examination procedures used from which pre-injection provocative clinical tests were identified.

Additional file 2: Diagnostic ultrasound and magnetic resonance arthrogram procedures. Description of the diagnostic ultrasound and magnetic resonance arthrogram imaging protocols used in this study.

Additional file 3: Subacromial bursa injection procedure. Video file showing ultrasound guided injection of local anaesthetic into the subacromial bursa.

Additional file 4: Acromioclavicular joint injection procedure. Video file showing injection of local anaesthetic into the acromioclavicular joint under fluoroscopic guidance.

Additional file 5: Glenohumeral joint injection procedure. Video file showing injection of local anaesthetic into the glenohumeral joint under fluoroscopic guidance.

Additional file 6: Association between x-ray and ultrasound variables and positive anaesthetic responses to subacromial bursa and acromioclavicular joint diagnostic blocks. Table showing additional results for x-ray and ultrasound imaging variables that were not associated with positive anaesthetic responses to subacromial bursa and acromioclavicular diagnostic blocks.

Additional file 7: Association between x-ray and ultrasound variables and positive anaesthetic responses to glenohumeral joint diagnostic block. Table showing additional results for x-ray and ultrasound imaging variables that were not associated with positive anaesthetic response to glenohumeral joint diagnostic block.

Additional file 8: Association between magnetic resonance arthrogram variables and positive anaesthetic responses to glenohumeral joint diagnostic block. Table showing additional results for magnetic resonance imaging variables that were not associated with positive anaesthetic response to glenohumeral diagnostic block.

Author details

¹Health & Rehabilitation Research Institute, AUT University, Akoranga Drive, Private Bag 92006, Northcote, Auckland, New Zealand. ²Physiosouth, Christchurch, New Zealand. ³Christchurch Radiology Group, Christchurch, New Zealand.

Authors' contributions

All authors were involved with conception and design of the study. AC performed all the clinical examination and pre-and post-injection clinical tests, collected and managed all data, carried out the preliminary analysis and drafted the manuscript. MC was involved in selection of guided diagnostic block and imaging procedures, performed and reported MRA procedures and provided radiological guidance in interpretation and discussion of results. ML, WH and PM contributed to methodological development, interpretation of data and critical appraisal of the manuscript for academic and clinical content. All authors read and approved the final manuscript.

Authors' information

This research was conducted as part of a larger diagnostic accuracy study which is the topic of AC's PhD thesis being conducted through AUT University, Auckland, New Zealand. Separate manuscripts are in preparation that will report the results of diagnostic accuracy calculations and the predictive ability of clinical examination and imaging findings to identify pain arising from specific structures, and specific shoulder pathology.

Competing interests

The authors declare that they have no competing interests.

References

1. van der Windt DAWM, Koes BW, De Jong BA, Bouter LM: **Shoulder disorders in general practice: Incidence, patient characteristics, and management.** *Ann Rheum Dis* 1995, **54(12)**:959-964.
2. Luime JJ, Koes BW, Hendriksen IJM, Burdorf A, Verhagen AP, Miedema HS, Verhaar JA: **Prevalence and incidence of shoulder pain in the general population: a systematic review.** *Scand J Rheumatol* 2004, **33(2)**:73-81.
3. Croft PR, Pope DP, Silman AJ: **The clinical course of shoulder pain: prospective cohort study in primary care.** *Br Med J* 1996, **313(7057)**:601-602.
4. Urwin M, Symmons D, Allison T, Brammah T, Busby H, Roxby M, Simmons A, Williams G: **Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation.** *Ann Rheum Dis* 1998, **57**:649-655.
5. van der Windt DAWM, Koes BW, Bocke AJP, Deville W, de Jong BA, Bounter LM: **Shoulder disorders in general practice: prognostic indicators of outcome.** *Br J Gen Pract* 1996, **46**:519-523.
6. Mitchell C, Adebajo A, Hay E, Carr A: **Shoulder pain: diagnosis and management in primary care.** *Br Med J* 2005, **331**:1124-1128.
7. Bot SDM, van der Waal JM, Terwee CB, Van der Windt DAWM, Schellevis FG, Bouter LM: **Incidence and prevalence of complaints of the neck and upper extremity in general practice.** *Ann Rheum Dis* 2005, **64(1)**:118-123.
8. Chakravarty KK, Webley M: **Disorders of the shoulder: an often unrecognised cause of disability in elderly people.** *Br Med J* 1990, **300**:848-849.
9. Chard MD, Hazelman R, Hazelman BL, King RH, Reiss BB: **Shoulder disorders in the elderly: a community survey.** *Arthritis Rheum* 1990, **34(6)**:766-769.
10. Feleus A, Bierma-Zeinstra SM, Miedema HS, Bernsen RM, Berhaar JA, Koes BW: **Incidence of non-traumatic complaints of arm, neck and shoulder in general practice.** *Man Ther* 2008, **13**:426-433.
11. Ostor AJK, Richards CA, Prevost AT, Speed CA, Hazleman BL: **Diagnosis and relation to general health of shoulder disorders presenting to primary care.** *Rheumatology* 2005, **44**:800-805.
12. Schellingerhout JM, Verhagen AP, Thomas S, Koes BW: **Lack of uniformity in diagnostic labeling of shoulder pain: time for a different approach.** *Man Ther* 2008, **13**:478-483.
13. Boocock MG, Collier JMK, McNair PJ, Simmonds M, Larmer PJ, Armstrong B: **A framework for the classification and diagnosis of work-related upper extremity conditions: systematic review.** *Semin Arthritis Rheum* 2009, **38(4)**:296-311.
14. Buchbinder R: **Classification of soft-tissue disorders of the neck and upper limb: do they satisfy methodological guidelines?** *J Clin Epidemiol* 1996, **49(2)**:141-149.
15. Hegedus EJ, Goode A, Campbell S, Morin A, Tamaddoni M, Moorman CT, Cook C: **Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests.** *Br J Sports Med* 2008, , **42**: 80-92.
16. Awerbuch MS: **The clinical utility of ultrasonography for rotator cuff disease, shoulder impingement syndrome and subacromial bursitis.** *Med J Aust* 2008, **188(1)**:50-53.
17. Shahabpour M, Kichouh M, Laridon E, Gielen JL, De Mey J: **The effectiveness of diagnostic imaging methods for the assessment of soft

tissue and articular disorders of the shoulder and elbow. *Eur Radiol* 2008, **65**(2):194-200.

18. Milgrom C, Schaffler M, Gilbert S, van Holsbeeck M: **Rotator-cuff changes in asymptomatic adults. The effect of age, hand dominance and gender.** *J Bone Joint Surg Br* 1995, **77**:296-298.

19. Connor PM, Banks DM, Tyson AB, Coumas JS, Alessandro DF: **Magnetic resonance imaging of the asymptomatic shoulder of overhead athletes.** *Am J Sports Med* 2003, **31**(5):724-727.

20. Miniaci A, Mascia AT, Salonen DC, Becker EJ: **Magnetic resonance imaging of the shoulder in asymptomatic professional baseball pitchers.** *Am J Sports Med* 2002, **30**(1):66-73.

21. Shubin 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**(3):204-208.

22. Reilly P, Macleod I, Macfarlane R, Windley J, Emery RJH: **Dead men and radiologists don't lie: A review of cadaveric and radiological studies of rotator cuff tear prevalence.** *Ann R Coll Surg Eng* 2006, **88**(2):116-121.

23. Shubin Stein BE, Ahmad CS, Pfaff CH, Bigliani LU, Levine WN: **A comparison of magnetic resonance imaging findings of the acromioclavicular joint in symptomatic versus asymptomatic patients.** *J Shoulder Elbow Surg* 2006, **15**(1):56-59.

24. Davidson M: **The interpretation of diagnostic tests: a primer for physiotherapists.** *Aust J Physiotherapy* 2002, **48**:227-233.

25. De Maeseneer M, Van Roy F, Lenchik L, Shahabpour M, Jacobson J, Ryu KN, Handelberg F, Osteaux M: **CT and MR arthrography of the normal and pathologic anterosuperior labrum and labral-bicipital complex.** *Radiographics* 2000, **20**(Suppl):67-81.

26. Sammarco VJ: **Os acromiale: frequency, anatomy, and clinical implications.** *J Bone Joint Surg Am* 2000, **82**(3):394-400.

27. Cyriax JH: **Textbook of Orthopaedic Medicine, Volume 1: Diagnosis of Soft Tissue Lesions.** Minneapolis: Balliere Tindall;, 8 1982.

28. Neer CS: **Impingement lesions.** *Clin Orthop* 1983, **173**:70-78.

29. Dreyfuss PH, Michaelsen M, Pauza K, McLarty J, Bogduk N: **The value of history and physical examination in diagnosing sacroiliac joint pain.** *Spine* 1996, **21**:2594-2602.

30. Strobel K, Pfirrmann CWA, Zanetti M, Nagy L, Hodler J: **MRI features of the acromioclavicular joint that predict pain relief from intraarticular injection.** *Am J Roentgenology* 2003, **181**:755-760.

31. Walton J, Mahajan S, Paxinos A: **A positive result on both the Paxinos test and bone scan ruled in a diagnosis of acromioclavicular joint pain.** *J Bone Joint Surg Am* 2004, **86**(A):807-812.

32. Anderson J, Read JW, Steinweg J: **The shoulder and arm.** In *Atlas of Imaging in Sports Medicine. Volume 1..* 1 edition. Edited by: Sydney. Australia: McGraw-Hill Book Company Australia Pty Limited; 1998:96-129.

33. McNally EG: **Practical Musculoskeletal Ultrasound.** London: Churchill Livingstone;, 1 2004.

34. Backhaus M, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA, Wakefield RJ, Manger B: **Guidelines for musculoskeletal ultrasound in rheumatology.** *Ann Rheum Dis* 2001, **60**(7):641-649.

35. Stoller DW, Wolf EM, Li AE, Nottage WM, Tirman PFJ: **The Shoulder.** In *Magnetic Resonance Imaging in Orthopaedics and Sports Medicine. Volume 2..* 3 edition. Edited by: Stoller DW. Baltimore: Lippincott Williams and Wilkins; 2007:1131-1462.

36. Flahault A, Cadilhac M, Thomas G: **Sample size calculation should be performed for design accuracy in diagnostic test studies.** *J Clin Epidemiol* 2005, **58**(8):859-862.

37. Bogduk N: **Practice guidelines for spinal diagnostic & treatment procedures.** San Francisco: International Spine Intervention Society;, 1 2004.

38. Speed CA, Hazleman BL: **Calcific tendinitis of the shoulder.** *N Engl J Med* 1999, **340**(20):1582-1584.

39. Farin PU, Jaroma H, Harju A, Soimakallio S: **Shoulder impingement syndrome: sonographic evaluation.** *Radiology* 1990, **176**(3):845-849.

40. Kolla S, Motamedi K: **Ultrasound evaluation of the shoulder.** *Semin Musculoskelet Radiol* 2007, **11**(2):117-125.

41. Naredo E, Aguado P, De Miguel E, Uson J, Mayordomo L, Gijon-Banos J, Martin-Mola E: **Painful shoulder: Comparison of physical examination and ultrasonographic findings.** *Ann Rheum Dis* 2002, **61**(2):132-136.

42. Van Holsbeeck M, Strouse PJ: **Sonography of the shoulder: Evaluation of the subacromial-subdeltoid bursa.** *Am J Roentgenology* 1993, **160**(3):561-564.

43. Tsai YH, Huang TJ, Hsu WH, Huang KC, Li YY, Peng KT, Hsu RWW: **Detection of subacromial bursa thickening by sonography in shoulder impingement syndrome.** *Chang Gung Med J* 2007, **30**(2):135-141.

44. Le Corroller T, Cohen M, Aswad R, Pauly V, Champsaur P: **Sonography of the painful shoulder: Role of the operator's experience.** *Skeletal Radiol* 2008, **37**(11):979-986.

45. Bruyn GAW, Naredo E, Moller I, Moragues C, Garrido J, De Bock GH, D'Agostino MA, Filippucci E, Lagnocco A, Backhaus M, Swen WAA, Balint P, Pineda C, Milutinovic S, Kane D, Kaeley G, Narvaez FJ, Wakefield RJ, Narvaez JA, De Augustin J, Schmidt WA: **Reliability of ultrasonography in detecting shoulder disease in patients with rheumatoid arthritis.** *Ann Rheum Dis* 2009, **68**(3):357-361.

46. Bruyn GAW, Pineda C, Hernandez-Diaz C, Ventura-Rios L, Moya C, Garrido J, Groen H, Pena A, Espinosa R, Moller I, Filippucci E, Iagnocco A, Balint PV, Kane D, D'Agostino MA, Angulo M, Ponte R, Fernandez-Gallardo JM, Naredo E: **Validity of ultrasonography and measures of adult shoulder function and reliability of ultrasonography in detecting shoulder synovitis in patients with rheumatoid arthritis using magnetic resonance imaging as a gold standard.** *Arthritis Care Res* 2010, **62**(8):1079-1086.

47. Naredo E, Moller I, Moragues C, De Agustin JJ, Scheel AK, Grassi W, De Miguel E, Backhaus M, Balint P, Bruyn GAW, D'Agostino MA, Filippucci E, Iagnocco A, Kane D, Koski JM, Mayordomo L, Schmidt WA, Swen WAA, Szkudlarek M, Terslev L, Torp-Pedersen S, Uson J, Wakefield RJ, Werner C: **Interobserver reliability in musculoskeletal ultrasonography: Results from a "Teach the Teachers" rheumatologist course.** *Ann Rheum Dis* 2006, **65**(1):14-19.

48. Sher JS, Uribe JW, Posada A, Murphy BJ, Zlatkin MB: **Abnormal findings on magnetic resonance images of asymptomatic shoulders.** *J Bone Joint Surg Am* 1995, **77**(1):10-15.

49. Yamanaka K, Matsumoto T: **The joint side tear of the rotator cuff. A followup study by arthrography.** *Clin Orthop Rel Res* 1994, **304**:68-73.

50. Ottenheijm RP, Jansen MJ, Staal JB, Van Den Bruel A, Weijers RE, De Bie RA, Dinant GJ: **Accuracy of diagnostic ultrasound in patients with suspected subacromial disorders: A systematic review and meta-analysis.** *Arch Phys Med Rehab* 2010, **91**(10):1616-1625.

51. Middleton WD, Teefey SA, Yamaguchi K: **Sonography of the rotator cuff: Analysis of interobserver variability.** *Am J Roentgenology* 2004, **183**(5):1465-1468.

52. Sher JS, Iannotti JP, Williams GR, Herzog RJ, Kneeland JB, Lisser S, Patel N: **The effect of shoulder magnetic resonance imaging on clinical decision making.** *J Shoulder Elbow Surg* 1998, **7**(3):205-209.

53. Alasaarela E, Takalo R, Tervonen O, Hakala M, Suramo I: **Sonography and MRI in the evaluation of painful arthritic shoulder.** *Br J Rheumatol* 1997, **36**(9):996-1000.

54. Ardic F, Kahraman Y, Kacar M, Kahraman MC, Findikoglu G, Yorgancioglu ZR: **Shoulder impingement syndrome: relationships between clinical, functional and radiologic findings.** *Am J Phys Med Rehabil* 2005, **85**(1):53-60.

55. Chang D, Mohana-Borges A, Borso M, Chung CB: **SLAP lesions: Anatomy, clinical presentation, MR imaging diagnosis and characterization.** *Eur Radiol* 2008, **68**(1):72-87.

56. Walch G, Boileau P, Noel E, Donell ST: **Impingement of the deep surface of the supraspinatus tendon on the posterosuperior glenoid rim: An arthroscopic study.** *J Shoulder Elbow Surg* 1992, **1**(5):238-245.

57. Pradhan RL, Itoi E: **Rotator interval lesions of the shoulder joint.** *Orthopedics* 2001, **24**(8):798-801.

58. Middleton WD, Reinus WR, Totty WG: **US of the biceps tendon apparatus.** *Radiology* 1985, **157**(1):211-215.

59. Iagnocco A, Coari G, Leone A, Valesini G: **Sonographic study of painful shoulder.** *Clin Exp Rheumatol* 2003, **21**(3):355-358.

60. Iagnocco A, Filippucci E, Meenagh G, Delle Sedie A, Riente L, Bombardieri S, Grassi W, Valesini G: **Ultrasound imaging for the rheumatologist. I ultrasonography of the shoulder.** *Clin Exp Rheumatol* 2006, **24**(1):6-11.

61. Lange U, Piegsa M, Teichmann J, Neeck G: **Ultrasonography of the glenohumeral joints: A helpful instrument in differentiation in elderly onset rheumatoid arthritis and polymyalgia rheumatica.** *Rheumatol Int* 2000, **19**(5):185-189.

62. Peat J, Barton B: **Categorical variables: risk statistics.** In *Medical Statistics: A guide to data analysis and critical appraisal..* 1 edition. Edited by: Banks M. Massachusetts: Blackwell Publishing Inc; 2005:241-266.

63. Arrol B, Robb G, Kool B, Astley T, Ball C, Carroll M, Durie R, Gendal P, Hanna C, Kofoed R, Mayhew J, Moran R, Reid D, Watt J: **The diagnosis and management of soft tissue shoulder injuries and related disorders: best practice evidence-based guideline.** Accident Compensation Corporation; 2004.

Pre-publication history

The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2474/12/119/prepub

Table of Self-Report Questionnaires and Clinical Examination Tests

Examination component	Response outcome	Thesis location	Reference
Self Report Questionnaires			
SF-8 Health Survey	%	Appendix 12, p282	(Ware et al., 2001)
Shoulder Pain and Disability Index (SPADI)	%	Appendix 13, p284	(Roach et al., 1991)
Fear Avoidance Beliefs Questionnaire (FABQ)	%	Appendix 14, p285	(Gordon Waddell et al., 1993)
Clinical examination			
Patient history			
Medical screening questionnaire	Dichotomous responses	Appendix 15, p287	
Symptom chart	Dichotomous responses	Appendix 16, p288	
Patient history	Dichotomous responses	Appendix 17, p290	
Physical examination		Appendix 18, p292	
Height/weight	Standard SI units		
Observation:	Present/absent		
Supraspinatus atrophy			
Infraspinatus atrophy			
ACJ swelling/thickening			
Cervical spine testing:	Yes/No	Appendix 18, p292	(Maitland, 1986)
Reproduction of shoulder symptoms			
Production of cervical spine symptoms			
Shoulder active ROM:		Appendix 18, p292	
[a] Elevation (flexion)	ROM (^{o}); symptom responses: Yes/No	Figure 3.3, p93	
Hand-behind-back	ROM (cm); symptom responses: Yes/No		
Hand-behind-head	ROM (cm); symptom responses: Yes/No		

[a] **Shoulder passive ROM:**	ROM (^{o}); symptom responses: Yes/No	Appendix 18, p292.	
Glenohumeral abduction		Figure 3.4a, p93	
External rotation (0° abd)		Figure 3.4b, p93	
External rotation (90° abd)		Figure 3.4c, p93	
Internal rotation (90° abd)		Figure 3.4d, p93	
Cross-body adduction (internal rotation)	Symptom responses: Yes/No		
[a] **Shoulder resisted tests:**	Peak muscle force (kg); symptom responses: Yes/No	Appendix 18, p292	
Abduction		Figure 3.5a, p94	
External rotation		Figure 3.5b, p94	
Internal rotation		Figure 3.5c, p94	
Scapuloclavicular tests:	Symptom responses: Yes/No	Appendix 18, p292	(Laslett, 1996)
Elevation			
Depression			
Protraction			
Retraction			
Orthopaedic special tests:	Positive/Negative according to original test criteria.	Appendix 18, p292	
Painful arc abduction			(Kessel & Watson, 1977)
Hawkins-Kennedy test			(Hawkins & Kennedy, 1980)
Empty can test			(F. Jobe & Moynes, 1982)
Drop-arm test			(Codman, 1934)
External rotation lag sign			(Hertel et al., 1996)
[a] Belly press test		Figure 3.5d, p94	(Barth et al., 2006)
Active compression test			(O'Brien et al., 1998)
Speed's test			(Gill et al., 2007)
Apprehension/Relocation test			(F. W. Jobe & Kvitne, 1989)
Palpation			(Mattingly & Mackarey, 1996)
Greater tuberosity (supraspinatus insertion)			
Lesser tuberosity (subscapularis insertion)			
Long head of biceps tendon			

Abbreviations. ACJ, acromioclavicular joint; ROM, range of motion

[a]three trials performed

Graph of Anaesthetic Responses Distributions Following Diagnostic Blocks
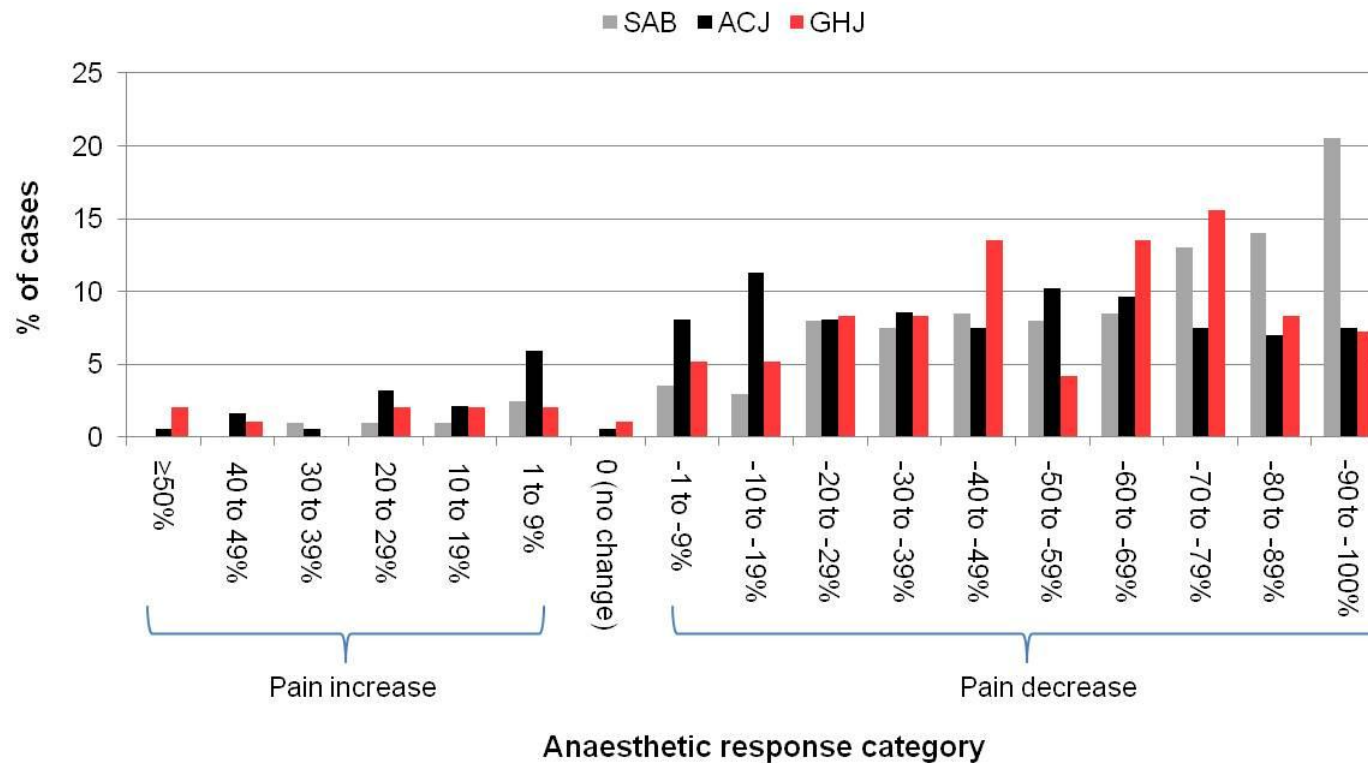
Distribution of anaesthetic responses following SAB, ACJ and GHJ diagnostic blocks. Graph shows the percent of participants in each anaesthetic response category, including the proportion reporting a post-injection increase or decrease in pain intensity. Negative (-) values indicate a reduction in pain intensity. SAB, subacromial bursa; ACJ, acromioclavicular joint; GHJ, glenohumeral joint.

Appendix 9

Table: Association Between X-ray and Ultrasound Findings and Response to SAB and ACJ Diagnostic Blocks

Pathology identified on imaging	SAB injection n=196 (PAR n=66)				ACJ injection n=153 (PAR n= 22)			
	% with pathology present reporting PAR	% with pathology absent reporting PAR	OR (95% CI)	Fishers test (*p*-value)	% with pathology present reporting PAR	% with pathology absent reporting PAR	OR (95% CI)	Fisher's test (*p*-value)
X-ray								
ACJ pathology	25	35	0.62 (0.26,1.47)	0.312	14	16	2.11 (0.69, 6.52)	0.189
arthropathy/degenerative changes	25	35	0.63 (0.24,1.68)	0.489	17	16	1.86 (0.55, 6.27)	0.296
osteolysis	14	34	0.32 (0.04, 2.74)	0.428	0	17	1.20 (0.13, 10.79)	1.000
Acromion type								
type I	37	31	1.28 (0.71, 2.34)	0.446	19	14	1.69 (0.68, 4.19)	0.352
type II	28	39	0.63 (0.35, 1.15)	0.132	15	18	0.59 (0.28, 1.49)	0.357
type III	75	33	6.24 (0.64, 61.23)	0.109	0	16	2.02 (0.20, 20.30)	0.469
type IV	75	33	0.66 (0.62, 0.73)	1.000	0	15	0.86 (0.80, 0.91)	1.000
os acromiale	75	33	6.14 (0.63, 60.25)	0.112	0	16	0.85 (0.80, 0.91)	1.000
Glenohumeral joint pathology	20	34	0.48 (0.10, 2.31)	0.500	0	17	0.85 (0.80, 0.91)	0.594
degenerative changes	13	35	0.27 (0.03, 2.25)	0.271	0	17	0.85 (0.80, 0.91)	0.594
Rotator cuff calcification	44	32	1.66 (0.71, 3.89)	0.262	13	16	0.30 (0.04, 2.36)	0.312
supraspinatus	56*	31	2.82 (1.00, 7.97)	0.054	25	16	0.58 (0.07, 4.74)	1.000
infraspinatus	29	34	0.79 (0.15, 4.21)	1.000	0	16	0.85 (0.80, 0.91)	0.594
subscapularis	0	34	0.66 (0.59, 0.73)	0.181	0	17	0.85 (0.80, 0.91)	0.594

Ultrasound

ACJ pathology	38	32	1.29 (0.66, 2.53)	0.489	11	16*	1.16 (0.22, 6.17)	0.053
Glenohumeral joint effusion	14	35	0.31 (0.04, 2.65)	0.426	20	15	0.85 (0.80, 0.91)	0.593
Rotator cuff - any pathology	38	30	1.44 (0.79, 2.62)	0.290	13	19	0.48 (0.19, 1.23)	0.166
any tear	42	31	1.60 (0.83, 3.06)	0.175	0	21*	0.32 (0.09, 1.15)	0.082
calcification	35	33	1.11 (0.56, 2.20)	0.861	18	16	0.72 (0.23, 2.28)	0.785
tendinosis	31	34	0.87 (0.37, 2.03)	0.834	25	14	0.59 (0.13, 2.73)	0.740
Supraspinatus pathology	39	30	1.50 (0.83, 2.72)	0.223	12	20	0.72 (0.28, 1.83)	0.643
calcification	49*	31	2.13 (1.00, 4.55)	0.068	22	15	1.40 (0.42, 4.60)	0.526
tendinosis	33	34	0.98 (0.42, 2.33)	1.000	27	14	0.63 (0.14, 2.92)	0.740
tear	40	32	1.43 (0.72, 2.85)	0.369	0	21	0.43 (0.12, 1.55)	0.286
intrasubstance	39	33	1.31 (0.53, 3.20)	0.640	0	18	0.32 (0.04, 2.53)	0.473
partial thickness (BS)	0	34	0 66 (0.59, 0.73)	0.302	0	17	0.85 (0.80, 0.91)	1.000
partial thickness (AS)	25	34	0.65 (0.13, 3.30)	0.720	0	17	2.08 (0.39, 11.05)	0.323
full thickness	70**	32	5.02 (1.25, 20.11)	0.033	0	16	0.84 (0.79, 0.91)	0.630
Infraspinatus pathology	33	34	0.98 (0.29, 3.40)	1.000	0	17	0.84 (0.79, 0.91)	0.217
calcification	33	34	0.98 (0.24, 4.07)	1.000	0	17	0.84 (0.79, 0.91)	0.359
tendinosis	0	34	0.66 (0.60, 0.73)	1.000	0	16	0.86 (0.80, 0.91)	1.000
tear	33	34	0.99 (0.09, 11.06)	1.000	0	16	0.85 (0.80, 0.91)	1.000
intrasubstance tear	100	33	0.33 (0.27, 0.41)	0.337	0	16	0.86 (0.80, 0.91)	1.000
partial thickness	0	34	0.66 (0.60, 0.73)	1.000	0	16	0.86 (0.80, 0.91)	1.000
full thickness	0	34	0.66 (0.60, 0.73)	1.000	0	16	0.86 (0.80, 0.91)	1.000
Subscapularis Pathology	28	35	0.72 (0.30, 1.72)	0.528	17	15	0.25 (0.03, 1.96)	0.203
calcification	25	35	0.63 (0.22, 1.81)	0.462	22	15	0.40 (0.05, 3.19)	0.698
tendinosis	0	34	0.66 (0.60, 0.73)	0.302	33	15	0.86 (0.80, 0.91)	1.000
tear	30	34	0.84 (0.21, 3.35)	1.000	0	17	0.85 (0.79, 0.91)	0.359
intrasubstance	40	34	1.32 (0.22, 8.12)	1.000	0	16	0.85 (0.80, 0.91)	1.000
partial thickness	25	34	0.65 (0.07, 6.39)	1.000	0	16	0.85 (0.80, 0.91)	1.000
full thickness	0	34	0.66 (0.60, 0.73)	1.000	0	16	0.85 (0.80, 0.91)	1.000

Long head of biceps tendon								
tendon sheath effusion	39	33	1.27 (0.54, 2.98)	0.657	46	10	0.59 (0.13, 2.73)	0.740
tendinosis	33	34	0.99 (0.09, 11.06)	1.000	0	16	3.07 (0.27, 35.39)	0.374
tear or rupture	33	34	0.99 (0.09, 11.06)	1.000	0	16	0.85 (0.80, 0.91))	1.000
subluxation	33	34	0.99 (0.09, 11.06)	1.000	0	16	0.85 (0.80, 0.91))	1.000
Subacromial bursa pathology								
bursal fluid/effusion	30	34	0.83 (0.30, 2.27)	0.807	0	18	0.84 (0.78, 0.90)	0.130
calcification	33	34	0.99 (0.09, 11.1)	1.000	0	15	1.17 (1.09, 1.25)	1.000
bursal dimension								
<1.0mm	34	33	1.01 (0.55, 1.87)	1.000	21	13	1.49 (0.60, 3.71)	0.480
≥1mm	33	34	0.99 (0.54, 1.84)	1.000	13	21	0.67 (0.27, 1.67)	0.480
≥2mm	36	33	1.16 (0.58, 2.30)	0.725	16	16	0.84 (0.29, 2.45)	1.000
≥3mm	40	34	1.32 (0.22, 8.12)	1.000	0	16	0.85 (0.80, 0.91)	1.000
bunching								
acromion	37	33	1.18 (0.64, 2.15)	0.645	20	12	0.71 (0.28, 1.81)	0.495
symptomatic bunching (acromion)	39	33	1.26 (0.68, 2.33)	0.528	19	13	0.57 (0.21, 1.56)	0.342
CAL	28	29	0.98 (0.38, 2.35)	1.000	5	14	1.21 (0.29, 4.92)	1.000
symptomatic bunching (CAL)	33	25	1.44 (0.58, 3.60)	0.488	7	12	1.25 (0.31, 5.10)	1.000

Abbreviations. SAB, subacromial bursa; ACJ, acromioclavicular joint; PAR, positive anaesthetic response (≥80% post-injection pain intensity reduction); OR, unadjusted odds ratio for PAR; CI, confidence interval; BS, bursal surface; AS, articular surface; CAL, coracoacromial ligament.

Note. Percentages do not total 100% as these represent proportion of participants with or without pathology on imaging (row percentages in contingency table) who experienced PAR. Negative anaesthetic response group results are not presented.

† no cases in which pathology was identified and OR could not be calculated.

** $p<0.05$. * $p<0.10$

Appendix 10

Table: Association Between X-ray and Ultrasound Findings and Response to GHJ Diagnostic Block

Pathology identified on imaging	% with pathology present reporting PAR	% with pathology absent reporting PAR	OR (95% CI)	Fisher's test (*p*-value)
GHJ injection n=73 (PAR n=12)				
X-ray (n=203)				
ACJ pathology	14	16	0.85 (0.16, 4.23)	1.000
arthropathy/degenerative changes	17	16	1.06 (0.19, 5.38)	1.000
osteolysis	0	17	0.83 (0.75, 0.92)	1.000
Acromion type				
type I	19	14	1.54 (0.45, 5.34)	0.534
type II	15	18	0.80 (0.23, 2.74)	0.759
type III	0	16	0.84 (0.76, 0.93)	1.000
type IV	0		0.84 (0.76, 0.93)	1.000
os acromiale	0	16	0.84 (0.76, 0.93)	1.000
Glenohumeral joint pathology	0	17	0.83 (0.75, 0.92)	1.000
degenerative changes	0	17	0.83 (0.75, 0.92)	1.000
Rotator cuff calcification	13	16	0.73 (0.08, 6.29)	1.000
supraspinatus	25	16	1.82 (0.17, 18.49)	0.510
infraspinatus	0	16	0.84 (0.76, 0.93)	1.000
subscapularis	0	17	0.83 (0.75, 0.92)	1.000
Ultrasound (n=203)				
ACJ pathology	11	16	0.60 (0.12, 3.07)	0.718
Glenohumeral joint effusion	20	15	1.45 (0.15, 14.34)	0.569
Rotator cuff - any pathology	13	19	0.61 (0.18, 2.13)	0.533
any tear	0	21	0.79 (0.69, 0.90)	0.029**
calcification	18	16	1.17 (0.28, 4.90)	1.000
tendinosis	25	14	2.0 (0.45, 8.83)	0.394
Supraspinatus pathology	12	20	0.55 (0.15, 2.03)	0.528
calcification	22	15	1.60 (0.29, 8.84)	0.611
tendinosis	27	14	2.29 (0.51, 10.30)	0.374
tear	0	21	0.79 (0.69, 0.90)	0.059*
intrasubstance	0	18	0.82 (0.73, 0.92)	0.338
partial thickness (BS)	0	17	0.83 (0.75, 0.92)	1.000
partial thickness (AS)	0	17	0.83 (0.75, 0.92)	0.583
full thickness	0	16	0.84 (0.76, 0.93)	1.000
Infraspinatus pathology	0	17	0.83 (0.75, 0.92)	1.000
calcification	0	17	0.83 (0.75, 0.92)	1.000
tendinosis	0	16	†	†
tear	0	16	†	†
intrasubstance tear	0	16	†	†
partial thickness	0	16	†	†
full thickness	0	16	†	†
Subscapularis Pathology	17	15	1.06 (0.20, 5.59)	1.000
calcification	22	15	1.60 (0.29, 8.84)	0.636
tendinosis	33	15	2.77 (0.23, 33.27)	0.421
tear	0	17	0.83 (0.75, 0.92)	1.000

intrasubstance	0	16	0.84 (0.76, 0.93)	1.000
partial thickness	0	16	0.84 (0.76, 0.93)	1.000
full thickness	0	16	0.84 (0.76, 0.93)	1.000
Long head of biceps tendon				
tendon sheath effusion	46	10	8.00 (2.02, 31.72)	0.004**
tendinosis	0	16	0.84 (0.76, 0.93)	1.000
tear or rupture	0	16	0.84 (0.76, 0.93)	1.000
subluxation	0	16	0.84 (0.76, 0.93)	1.000
Subacromial bursa pathology				
bursal fluid/effusion	0	18	0.82 (0.74, 0.92)	0.589
calcification	0	16	†	†
bursal dimension				
<1.0mm	21	13	1.86 (0.54, 6.47)	0.519
≥1mm	13	21	0.54 (0.16, 1.86)	0.524
≥2mm	16	16	0.90 (0.22, 3.73)	1.000
≥3mm	0	16	†	†
bunching				
acromion	20	12	1.85 (0.51, 6.74)	0.328
symptomatic bunching (acromion)	19	13	1.55 (0.42, 5.68)	0.496
CAL	5	14	0.35 (0.03, 3.70)	0.610
symptomatic bunching (CAL)	7	12	0.52 (0.05, 5.55)	1.000

Abbreviations: GHJ, glenohumeral joint; PAR, positive anaesthetic response (≥80% post-injection pain intensity reduction); CAL, coracoacromial ligament; OR, unadjusted odds ratio for PAR; CI, confidence interval; BS, bursal surface; AS, articular surface.

Percentages do not total 100% as these represent proportion of participants with or without pathology on imaging (row percentages in contingency table) who experienced PAR. Negative anaesthetic response group results are not presented.

† no cases in which pathology was identified and OR could not be calculated.
**significant at $p<0.05$. $p<0.10$

Appendix 11

Table: Association Between MRI Findings and Response to GHJ Diagnostic Block

Pathology identified on MRA	GHJ injection n=73 (PAR n=12)			
	% with PAR when pathology present	% with PAR when pathology absent	OR (95% CI)	Fisher's test (*p*-value)
ACJ pathology	20	11	2.03 (0.50, 8.23)	0.516
arthropathy	14	19	0.72 (0.21, 2.49)	0.751
osteolysis	40	15	3.93 (0.58, 26.58)	0.187
os acromiale				
GHJ pathology	14	21	0.59 (0.17, 2.05)	0.524
degenerative changes	0	18	†	0.591
labral tear	12	21	0.53 (0.15, 1.96)	0.360
paralabral cyst	22	16	1.57 (0.28, 8.69)	0.636
rotator interval pathology	15	21	0.59 (0.17, 2.06)	0.574
synovitis- inferior recess	22	16	1.54 (0.28, 8,53)	0.639
synovitis– rotator interval	14	18	0.69 (0.19, 2.55)	0.752
Rotator cuff - any pathology	11	27	0.34 (0.10, 1.21)	0.101
any tear	9	23	0.33 (0.08, 1.35)	0.124
tendinosis	9	20	0.42 (0.08, 2.10)	0.326
Supraspinatus pathology	12	22	0.52 (0.15, 1.81)	0.345
tendinosis	11	18	0.58 (0.11, 2.91)	0.718
tear	11	20	0.49 (0.12, 2.00)	0.350
intrasubstance	0	19	0.81 (0.72, 0.91)	0.198
partial thickness – BS	0	18	0.83 (0.74, 0.92)	0.583
partial thickness – AS	22	16	1.57 (0.28, 8.69)	0.636
full thickness	25	16	1.79 (0.17, 18.80)	0.521
Infraspinatus pathology	0	19	0.82 (0.72, 0.91)	0.198
tendinosis	0	17	0.83 (0.75, 0.92)	1.000
tear	0	18	0.82 (0.74, 0.92)	0.581
intrasubstance tear	0	17	0.83 (0.75, 0.92)	1.000
partial thickness	0	17	0.83 (0.75, 0.92)	1.000
full thickness		16	†	†
Subscapularis Pathology	0	19	0.82 (0.73, 0.92)	0.339
tendinosis	0	17	0.83 (0.75, 0.92)	1.000
tear	0	18	0.83 (0.74, 0.92)	0.583
intrasubstance	0	17	0.83 (0.75, 0.92)	1.000
partial thickness		16	†	†
full thickness	0	17	0.84 (0.76, 0.93)	1.000
Long head of biceps tendon				0.583
tendinosis	0	17	0.83 (0.75, 0.92)	1.000
tear or rupture	0	17	0.83 (0.75, 0.92)	1.000
Subacromial bursa				0.725
bursal thickening	20	15	1.47 (0.27, 8.08)	0.648
bursitis	15	20	0.71 (0.19, 2.69)	0.725
contrast in bursa on MRI	33	15	2.90 (0.47, 17.99)	0.254

Abbreviations. GHJ, glenohumeral joint; PAR, positive anaesthetic response (≥80% post-injection pain intensity reduction); MRA, magnetic resonance arthrogram; ACJ. acromioclavicular joint; BS, bursal surface; AS, articular surface; MRI, magnetic resonance imaging

† no cases in which pathology was identified- association could not be calculated.

** *p*<0.05. * *p*<0.10

Appendix 12

SF-8^{TM} General Health Survey.

The self-administered SF-8^{TM} questionnaire is a shortened form of the SF-36^{TM} Health Survey, a generic, multipurpose survey of health status, using a single question to measure each of the eight health domains included in the SF-36^{TM}. These include physical functioning, role limitations due to physical health problems, bodily pain, general health, vitality, social functioning, role limitations due to emotional problems and mental health. There are also two summary scores for "physical components" and "mental components" of the survey. The same scoring system is used as the SF36, enabling comparisons of population normative values with studies focused on specific populations. This validated shortened version of the SF-36^{TM} was selected to reduce the potential for 'questionnaire fatigue' due to the number of questionnaires included in this study. The 4-week recall version of the SF-8^{TM} survey was selected as the 24-hour and 1-week recall versions had not been validated at the time this study commenced and responsiveness was unknown (Ware et al., 2001).

Participants were reminded about the 4-week recall requirement, and scoring was carried out using the SF-8^{TM} online scoring system provided by QualityMetric, Inc (http://www.sf-36.org/demos/SF-8.html).

SF-8 Health Survey

 AUT UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

This survey asks for your views about your health. Answer every question by selecting the answer as indicated. If you are unsure about how to answer a question, please give the best answer you can.

For each of the following questions, please mark an [x] in the one box that best describes your answer.

1. Overall, how would you rate your health during the **past 4 weeks**.

Excellent	Very good	Good	Fair	Poor	Very poor
☐	☐	☐	☐	☐	☐

Office use: PF

2. During the **past 4 weeks**, how much did physical health problems limit your usual physical activities (such as walking and climbing stairs)?

Not at all	Very little	Somewhat	Quite a lot	Could not do physical activities
☐	☐	☐	☐	☐

RP

3. How much bodily pain have you had during the **past 4 weeks**?

None	Very mild	Mild	Moderate	Severe	Very Severe
☐	☐	☐	☐	☐	☐

BP

4. During the **past 4 weeks**, how much difficulty did you have doing your daily work, both at home and away from home, because of your physical health?

None at all	A little bit	Some	Quite a lot	Could not do daily work
☐	☐	☐	☐	☐

GH

5. During the **past 4 weeks**, how much energy did you have?

Very much	Quite a lot	Some	A little	None
☐	☐	☐	☐	☐

VT

6. During the **past 4 weeks**, how much did your physical health or emotional problems limit your usual social activities with family or friends?

Not at all	Very little	Somewhat	Quite a lot	Could not do social activities
☐	☐	☐	☐	☐

SF

7. During the **past 4 weeks**, how much have you been bothered by **emotional problems** (such as feeling anxious, depressed or irritable)?

Not at all	Slightly	Moderately	Quite a lot	Extremely
☐	☐	☐	☐	☐

RE

8. During the **past 4 weeks**, how much did personal or emotional problems keep you from doing your usual work, school or other daily activities?

Not at all	Very little	Somewhat	Quite a lot	Could not do daily activities
☐	☐	☐	☐	☐

MH

Appendix 13

Shoulder Pain and Disability Index (SPADI)

The SPADI (Appendix 13, p284) is designed to measure the pain and disability associated with shoulder pathology. The SPADI is a self-administered index consisting of 13 items divided into two subscales: pain and disability. Each question each uses a 10-point visual analogue scale (VAS) with anchors including 'no pain' and 'worst imaginable pain' (pain subscale), and 'no difficulty' and 'so difficult requires help' (disability subscale). The SPADI has demonstrated good internal consistency, test-retest reliability, and criterion and construct validity (J. C. MacDermid, Solomon, & Prkachin, 2006; Roach et al., 1991).

In the current study the questionnaire was explained to the participant in a standardised manner, after which they completed the questionnaire independently. Each subscale (pain and disability) scored as a percent of the total for each subscale, and as a % of the total of both subscales.

PAIN SCALE	
How severe is your pain:	
1. At its worst.	No pain 0 1 2 3 4 5 6 7 8 9 10 Worst Pain Imaginable
2. When lying on involved side.	No pain 0 1 2 3 4 5 6 7 8 9 10 Worst Pain Imaginable
3. Reaching for something on a high shelf.	No pain 0 1 2 3 4 5 6 7 8 9 10 Worst Pain Imaginable
4. Touching the back of your neck.	No pain 0 1 2 3 4 5 6 7 8 9 10 Worst Pain Imaginable
5. Pushing with the involved arm.	No pain 0 1 2 3 4 5 6 7 8 9 10 Worst Pain Imaginable
DISABILITY SCALE	
How much difficulty did you have:	
1. Washing your hair.	No difficulty 0 1 2 3 4 5 6 7 8 9 10 So difficult required help
2. Washing your back.	No difficulty 0 1 2 3 4 5 6 7 8 9 10 So difficult required help
3. Putting on an undershirt or pullover sweater.	No difficulty 0 1 2 3 4 5 6 7 8 9 10 So difficult required help
4. Putting on a shirt that buttons down the front.	No difficulty 0 1 2 3 4 5 6 7 8 9 10 So difficult required help
5. Putting on your pants.	No difficulty 0 1 2 3 4 5 6 7 8 9 10 So difficult required help
6. Placing an object on a high shelf.	No difficulty 0 1 2 3 4 5 6 7 8 9 10 So difficult required help
7. Carrying a heavy object of 10 pounds.	No difficulty 0 1 2 3 4 5 6 7 8 9 10 So difficult required help
8. Removing something from your back pocket.	No difficulty 0 1 2 3 4 5 6 7 8 9 10 So difficult required help

DEVELOPED BY Roach 1991 [1];

Reference List

1. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. Arthritis Care Res. 4[4], 143-149. 1991.

Appendix 14

Fear Avoidance Beliefs Questionnaire (FABQ)

The Fear Avoidance Beliefs Questionnaire (FABQ) (Appendix 14, p285) was originally developed by Waddell et al., (1993) in patients with chronic low back pain (Gordon Waddell et al., 1993). He identified fear avoidance beliefs about work, and about physical activity as two principal components of disability in activities of daily living and work loss due to low back pain. Since then, fear of pain has also been associated with baseline shoulder function (Lentz, Barabas, Day, Bishop, & George, 2009). When modified to replace "back pain" with "shoulder pain" the FABQ questionnaire has demonstrated substantial test-retest reliability, a high correlation with both SPADI pain and disability scores, and was a better than chance predictor of missing days off work during the 48 hour study period (Mintken, Cleland, Whitman, & George, 2010).

Participants in the study were instructed to complete the FABQ, emphasising what they "believed" about their pain, and percentages were calculated for each subscale (work, physical activity and general)

.

FABQ Questionnaire

		completely disagree			unsure			completely agree
1.	My pain is caused by physical activity	0	1	2	3	4	5	6
2.	*Physical activity makes my pain worse*	0	1	2	3	4	5	6
3.	*Physical activity may harm my shoulder*	0	1	2	3	4	5	6
4.	*I should not do physical activities that (might) make my pain worse.*	0	1	2	3	4	5	6
5.	*I cannot do physical activities which (might) make my pain worse.*	0	1	2	3	4	5	6
6.	**My pain was caused by my work or by an accident at work**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
7.	**My work aggravated my pain**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
8.	I have a claim for compensation for my pain	0	1	2	3	4	5	6
9.	**My work is too heavy for me**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
10.	**My work makes or would make my pain worse**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
11.	**My work might harm my shoulder**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
12.	**I should not do my normal work with my present pain**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
13.	I cannot do my normal work with my present pain	0	1	2	3	4	5	6
14.	I cannot do my normal work till my pain is treated	0	1	2	3	4	5	6
15.	**I do not think that I will be back to my normal work within 3 months**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
16.	I do not think that I will ever be able to go back to that work.	0	1	2	3	4	5	6

Scoring

FABQ - work _____ / 42

FABQ - physical activity _____ / 24

FABQ - general _____ / 30

Total _____ / 96

Appendix 15

Medical History

Medical Screening Questionnaire	AUT UNIVERSITY TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Exclusion Criteria:

Please answer "yes" or "no" to the following questions: (Please tick (☑) the appropriate box).

Yes	**No**	
☐	☐	Are you pregnant?
☐	☐	Are you currently breastfeeding?
☐	☐	Are you taking any blood thinning medication (anticoagulants eg warfarin, heparin)?
☐	☐	Have you previously had an adverse reaction (or allergy) to local anaesthetic Medication or contrast agents used during radiology procedures?
☐	☐	Have you ever had surgery to your neck or your painful shoulder?
☐	☐	Do you have any numbness or tingling anywhere in your arm, forearm or hand?
☐	☐	Do you have any metal implants in your body? (eg, rods, screws, clips, stents).
☐	☐	Do you have a pacemaker

If you have answered "Yes" to any of these questions, you may not be eligible to participate in this study. Please inform the researcher/examiner before completing the rest of this form

Medications:

Please list ALL medications you are currently taking:

General Health:

Do you suffer from any of the following? (Please tick (☑) the appropriate box).

Yes	**No**	
☐	☐	Asthma
☐	☐	Diabetes (insulin dependent Type 1)
☐	☐	Heart disease
☐	☐	Heart Arhythmias
☐	☐	High blood pressure
☐	☐	Liver disease/cirrhosis
☐	☐	Kidney disease
☐	☐	Epilepsy
☐	☐	Any other condition for which you see your doctor regularly?
		Please describe: _____
☐	☐	Have you experienced any recent, unexplained weight loss?

Smoking History:

Cigarettes per day:

☐ Non-smoker   ☐ 1-10   ☐ 11-20   ☐ 21-30   ☐ 30+

Surgical History:

Please list any surgery for which you had a general anaesthetic:

Date: **Surgery:**

__/__/__ _   _____

__/__/__   _____

Allergies:

Please list all known allergies:

_____

Appendix 16

Symptom Chart

Symptom Chart

AUT
UNIVERSITY
TE WĀNANGA ARONUI O TĀMAKI MAKAU RAU

Location of Symptoms:

1. Please indicate with a cross (X) the location of the main <u>primary or predominant area</u> of shoulder pain/symptoms.

2. Please shade in any <u>other areas</u> where your shoulder pain/symptoms are felt.

left   right   right   left
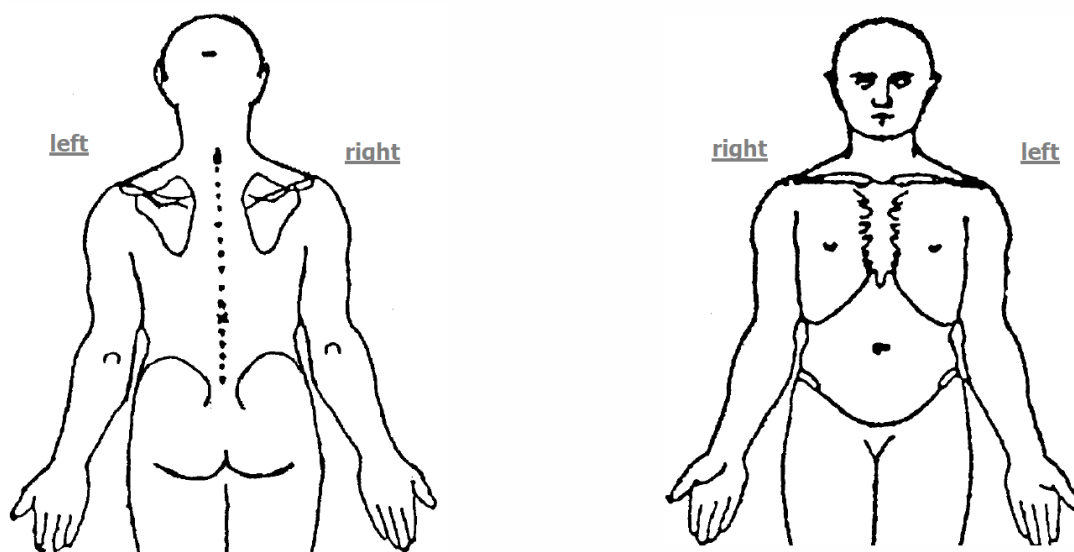
3. How would you describe your MAIN pain? ☐ Sharp pain   ☐ Aching pain   ☐Other: _____

4. Does your pain ever radiate to below your elbow?   ☐Yes   ☐No

Pain Location: ☐ Anterior	☐ Superior	☐ Lateral Sh	☐ Lateral Arm	☐ Posterior	☐ Axillary
☐ Chest	☐ Clavicle	☐ Elbow/forearm/hand		☐ CERVICAL SPINE SYMPTOMS	

Pain Intensity:

Visual Analogue Scale
Please mark the horizontal line below with a cross (X) to indicate how you would rate the intensity of your **most dominant shoulder pain** (the pain that you marked with an arrow (**X**) above)

5. The **BEST/LOWEST** level of pain you have experienced at ANY time in the last **48 hours**

0 ────────────────────────── **10**
No Pain   *Worst Imaginable Pain*

6. The **AVERAGE** amount of pain you have felt over the last **48 hours**

0 ────────────────────────── **10**
No Pain   *Worst Imaginable Pain*

7. The **WORST/HIGHEST** level of pain you have experienced at ANY time over the last **48 hours**

0 ────────────────────────── **10**
No Pain   *Worst Imaginable Pain*

8. Do you experience any of the following symptoms?:

☐ Yes ☐ No crepitus (grinding/graunching)
☐ Yes ☐ No clicking/clunking
☐ Yes ☐ No feelings of instability/"popping out"
☐ Yes ☐ No tingling/numbness anywhere in the arm/hand
☐ Yes ☐ No Other *(describe)*: _____

9. Is there a time of day/position of the arm when you are totally free of pain?

☐ Yes If "**Yes**" – go to Question 12
☐ No If "**No**" – go to Question 11

☐ Constant ☐ Intermittent

10. What effect does <u>resting the arm</u> have on your shoulder pain:

☐ Rest has **no effect** on my symptoms
☐ Rest makes my symptoms **worse**
☐ Rest **relieves** my symptoms

24 hour Pain Behaviour:

Morning

11. On rising/first thing in the morning my shoulder pain (compared to my "average" pain) is:

☐ Worse
☐ No different
☐ Better

12. Is your shoulder stiff first thing in the morning? ☐ Yes ☐ No

13. If "yes" How long until the shoulder stiffness resolves? _____ (hrs) _____ (mins)

During the Night

14. Does your shoulder pain wake you after you have gone to sleep? ☐ Yes ☐ No
15. Are you able to sleep on the side of your sore shoulder at night? ☐ Yes ☐ No

Aggravating Positions and Activities:

16. Do the following positions/activities produce or aggravate your shoulder pain?

Please place a 'tick' ☑ in the appropriate box:

Worst	Activity	Yes	No	Don't know
	Reaching up - above shoulder height /overhead activities provokes my shoulder pain			
	Movement in any direction provokes my shoulder pain			
	Rest/not moving my arm increases my shoulder pain			
	Sudden movements provoke my shoulder pain			
	Other *(please describe)*:			

17. Please indicate which of the above is the **ONE WORST** position/activity by placing a tick in the box next to that activity in the column labelled "WORST" (tick ONLY ONE).

Appendix 17

Patient History

Patient History	AUT

ONSET of SYMPTOMS

1. The <u>current</u> shoulder symptoms have been present since: _____/_____/_____ (date)

   (If you cannot remember the exact date, please give the closest estimate for the day/month/year)

2. **Describe how your shoulder symptoms started:**

   ☐ Trauma   ☐ Strain/Injury   ☐ Repetitive/Overuse   ☐ Unknown
   eg, fall, sudden   eg, lift heavy object,
   load or impact,   over-reach or over-stretch,
   external blow   throwing type activity

   Describe how the trauma/injury occurred, or describe the activity that caused your shoulder pain:

   _____

3. **When did you FIRST notice the pain?**

   ☐ Immediately   ☐ Within 48 hours   ☐ After more than 48 hours

   If the pain was not immediate, describe what you were doing when you FIRST noticed the pain:

   _____

4. **Would you say your shoulder was 100% before this happened?**   ☐ Yes   ☐ No

5. **From the time your symptoms FIRST started, when did you FIRST consult a medical professional** (Dr or physio etc) **about the problem?**

   ☐ within   ☐ within   ☐ within   ☐ within   ☐ within   ☐ more than
   1 week   1 month   3 months   6 months   12 months   12 months

PAST HISTORY of SHOULDER PROBLEMS (these questions relate only to <u>PAST</u> episodes of shoulder pain).

6. **Have you had previous problems with this (same) shoulder?**   ☐ Yes   ☐ No

7. **When was the FIRST episode of this shoulder problem?** ____/____/____ *(date)*

8. **How did the problem <u>first</u> start?** ☐ Trauma/Injury   ☐ Repetitive/Overuse   ☐ Unknown

   Describe: _____

9. **Did the problem <u>fully</u> resolve before the current episode?**   ☐ Yes   ☐ No

10. **Have you ever had problems with shoulder pain in the opposite shoulder?** ☐ Yes   ☐ No

FAMILY HISTORY of SHOULDER PROBLEMS

11. **Has anyone in your immediate family had a history of shoulder pain/problems**

    ☐ Yes   ☐ No   ☐ Don't know   Relationship of this person to you _____

    Did they require surgery for this problem?   ☐ Yes   ☐ No   ☐ Don't know

OCCUPATION DETAILS

12. **Are you currently in paid employment?**
    ☐ Yes   Occupation: _____ *- Go to Question 13.*
    ☐ No   *Please select from the following list:*
              ☐ Retired
              ☐ At home with children/maternity leave

13. Please indicate the demands of your work:

☐ Low shoulder demand   (eg, clerical worker)

☐ Moderate shoulder demand (eg, tradesperson)

☐ High shoulder demand   (eg, heavy lifting or frequent overhead work).

14. How is your shoulder pain affecting your ability to work?

☐ I can continue to work my normal duties and hours

☐ I am on light/restricted/different duties or reduced hours due to shoulder pain

☐ I have been off work since ___/___/___ due to my shoulder pain.

(insert date)

SPORT/RECREATION/LEISURE DETAILS

15.   Are you involved in any regular sport, recreational activities or hobbies?   ☐ Yes   ☐ No

Describe/list: _____

16. Please indicate the demands of your sport/recreational activity:

☐ Low shoulder demand   (eg, walking, running, hiking, lawn bowls, easy gardening, handcrafts)

☐ Moderate shoulder demand (eg, golf, fishing, moderate gardening, soccer, mountain biking)

☐ High shoulder demand   (eg, swimming, racket sports, overhead sports eg, water polo, volleyball, contact sport, heavy landscaping, throwing sports, weight lifting).

17.   Is your shoulder pain preventing you from participating in any of these activities?

☐ Yes   which ones?_____

☐ No

Appendix 18

Physical Examination

Physical Examination		AUT UNIVERSITY

Height: _____ cm   **Waist Circumference:** _____ cm

Weight: _____ kg   **Affected Shoulder:** R / L

6% BW: _____ kg   **Bilateral Symptoms:** Yes / No

STANDING

Observation

ATROPHY:	Yes	No	Unsure	Contralateral side affected	Comment
Supraspinatus					
Infraspinatus/T.Minor					
Biceps Belly					

OTHER JOINTS:	Yes	No	Unsure	Contralateral side affected	Comment
AC Joint swollen					
AC Joint thickened					
SC Joint swollen					
SC Joint thickened					

PAINFUL ARC ABDUCTION	Positive			Negative	Indeterminate	Comment	PPPT
	Raise only	Lower only	Raise AND Lower				
Painful Arc							

Patient has any physical limitations of the elbow, wrist or thumb?   ☐ No   ☐ Yes _____

ACTIVE ROM ROTATION	ROM (cm)		Symptoms Produced			Comment	PPPT
	Unaff	Affected	Yes	No	Unsure		
Hand Behind Back							
Hand Behind Head							
Difference (cm)							

ORTHOPAEDIC SPECIAL TESTS	Positive	Negative	Indeterm-inate	Comments	PPPT
Drop Arm Sign					
Hawkins Kennedy					
Empty Can/Jobe Test: *Pain*					
Weakness					
Active Compression Test *- ACJ*					
Labrum (pain)					
Labrum (click)					

SITTING

CERVICAL SPINE	Motion Restriction			Cervical Spine Symptoms Produced		Shoulder Symptoms Produced		Pain Location	PPPT
	Major	Minor	None	Yes	No	Yes	No		
Flexion									
Extension									
L) Rotation									
R) Rotation									
L) Lateral Flexion									
R) Lateral Flexion									

SCAPULOCLAVICULAR TESTS	Symptoms Produced			Indeterminate	Comment	PPPT
	Yes	No	Unsure			
Elevation						
Depression						
Protraction						
Retraction						

ACTIVE ROM ELEVATION	Symptoms Produced			ROM (Degrees)						Range Where Symptoms Produced				PPPT
	Yes	No	Unsure	Unaffected			Affected			Arc	Thru-range	End range	Over-pressure	
				T1	T2	T3	T1	T2	T3					
Flexion														

RESISTED MUSCLE TESTS	Symptoms Produced			Force						PPPT
	Yes	No	Unsure	Unaffected			Affected			
				T1	T2	T3	T1	T2	T3	
Abduction										
External Rotation										
Internal Rotation										
Elbow Flexion										
Biceps (Speeds test)										

PASSIVE ROM GLENOHUMERAL ABDUCTION	Symptoms Produced			ROM (degrees)						Pain LImitation (stop before force limit reached)		PPPT
	Yes	No	Unsure	Unaffected			Affected			Yes	No	
				T1	T2	T3	T1	T2	T3			
Glenohumeral Abduction												

ORTHOPAEDIC SPECIAL TESTS	Positive	Negative	Indetermi nate	Comment								PPPT
External Rotation Lag Sign												
Belly Press Test				Strength							Comment	
Pain				Unaffected Side			Affected Side					
Elbow Extension				T1	T2	T3	T1	T2	T3			
Shoulder forward												

PALPATION	Symptoms Produced				Comment
	Typical	Atypical	None	Indetermin ate	
ACJ					
SCJ					
Greater Tuberosity – supraspinatus insertion					
Greater Tuberosity – infraspinatus/T.min insertion					
Lesser Tuberosity					
Biceps Tendon					

SUPINE

PASSIVE ROM	Symptoms Produced			ROM (degrees)						Pain Limitation (stop before force limit reached)		PPPT
				Unaffected			Affected					
	Yes	No	Unsure	T1	T2	T3	T1	T2	T3	Yes	No	
External Rotation (0° Abd)												
External Rotation (90° Abd)												
Internal Rotation (90° Abd)												
TOTAL GHJ Rotation	███	███	███							███	███	███
Horizontal Adduction (Internal Rotation)												
Horizontal Adduction (External Rotation)												

ORTHOPAEDIC SPECIAL TESTS	Positive	Negative	Indeterminate	Comments	PPPT
Biceps Load II Test					
Apprehension test - Pain					
- Apprehension					
Relocation test - Pain					
- Apprehension					

Appendix 19

Calculation of Diagnostic Accuracy Statistics

Diagnostic accuracy statistics including sensitivity, specificity, positive (PPV) and negative predictive values (NPV), and positive (+LR) and negative (-LR) likelihood ratios were calculated using Confidence Interval Analysis software (Bryant, 2000). Post-test probabilities were also assessed for a positive and negative 'outcome' status. Statistics were calculated as shown below.

		Outcome status		
		Outcome present	Outcome absent	
Test result	Test positive	a	b	PPV=a/(a+b)
	Test negative	c	d	NPV=d/(c+d)
		Sensitivity=a/(a+c)	Specificity=d/(b+d)	

+LR=sensitivity/(1-specificity
-LR=(1-sensitivity)/specificity
Pre-test probability=a+c/a+b+c+d
Pre-test odds=(a+c)/1-(a+c)
Post-test odds (positive)=pre-test odds*+LR
Post-test probability=post-test odds/(post-test odds+1)

Calculation of Diagnostic Accuracy Statistics. Figure showing contingency cells and formulas for calculation of diagnostic accuracy statistics. Abbreviations: PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio.

Interpretation of Diagnostic Accuracy Statistics.

Sensitivity and specificity.

Sensitivity (probability of a positive test when the condition is present) and specificity (probability of a negative test when the condition is absent) are subject to interpretation according to the relative clinical importance of being able to rule-in or rule-out a particular condition. The higher the probability, the less chance there is of obtaining a false-negative or false-positive result.

Likelihood ratios.

Guidelines are available for the interpretation of likelihood ratios (Jaeschke, Guyatt, & Sackett, 1994), and these are used for interpretation of positive (+LR) and negative (-LR) likelihood ratios in this thesis. A +LR is defined as the increase in odds favouring the condition being present when the test is positive, and -LR is defined as the change in odds favouring the condition being present when the test is negative:

- Likelihood ratios greater than 10.0 or less than 0.1 generate large and often conclusive changes from pre-test to post-test probability.
- Likelihood ratios of 5.0 to 10.0, and 0.1 to 0.2 generate moderate shifts in pre-test to post-test probability.
- Likelihood ratios of 2.0 to 5.0, and 0.5 to 0.2 generate small (but sometimes important) changes in pre-test to post-test probability.
- Likelihood ratios of 1.0 to 2.0, and 0.5 to 0.1 alter probability to a small (and rarely important) degree.

Positive and negative predictive values and post-test probabilities.

Predictive values indicate how likely someone with a positive test is to have the disorder (positive predictive value (PPV)) and how likely it is that someone with a negative test does not have the disorder (negative predictive value (NPV)). The PPV also provides an indication of the 'post-test probability' of a condition being present, and can be compared to the pre-test probability of the condition being present (prevalence of the condition in the population under observation) to assess the value of the test in improving the ability to predict the presence or absence of the condition. Because predictive values are closely related to the prevalence of the condition (pre-test probability), they can only be generalised to settings in which the prevalence of the condition closely resembles that in the study (Jaeschke et al., 1994).