Master of Philosophy (MPhil)

Shoulder Pain Mapping
For Common Shoulder Disorders

University of Salford, Manchester

Health, Sport & Rehabilitation sciences

Levent Bayam

Student ID: @00279526

i) The full title of the thesis: Shoulder Pain Mapping For Common Shoulder Disorders

ii) The total number of volumes: 1

iii) The full name of the author: Levent Bayam

iv) The qualification for which the thesis is submitted: Master of Philosophy (MPhil)

v) The name of the University: University of Salford, Manchester

vi) The name of the School: Health, Sport & Rehabilitation sciences

vii) The calendar year of submission: 2013

Contents **page**

List of the Tables 7

List of the Figures 9

Acknowledgement 10

Declaration 11

Common Abbreviations 12

Abstract 13

Introduction 17

Nature of Literature Review 19

Aim of the Literature Search 19

Terms 20

Databases Search 21

Diagnosis of Common Shoulder Disorders 22

Classification of Common Shoulder Disorders and Their Descriptions 26

Importance of Mapping to inform diagnosis 30

Existing Literature in Pain Mapping 33

Identification of Necessity for this study 37

Method 39

Phase one 40

Method 40

Design 41

Participants 44

Inclusion Criteria 44

Exclusion Criteria 44

Recruitment 45
Procedure 45
Data collection 46
Data analysis (and analytic strategy) 47
Ethical implications 50
In summary- Phase one 51
Phase Two 53
Method 53
Design 53
Participants 54
Recruitment 54
Procedure 55
Data analysis 56
Ethical implications 57
In summary- Phase two 57
Phase Three 58
Method 58
Design 58
Participants 59
Procedure 59
Data analysis 60
Ethical implications 61
In summary- Phase three 61
Key Statistical Terminology 63
Results 65

Phase One 66

Character and Radiation of Pain for Common Shoulder Disorders 69

The Distribution Area of Pain and the Severity (First Phase Study) 70

Phase Two 72

Testing Colour-coded Shoulder Pain Maps 75

Additional Changes to Phase One Results in Distribution and Character of Pain in Shoulder Disorders 82

Classification of Common Shoulder Disorders According to Pain Distribution / Radiation 84

Shoulder Pain Mapping Algorithm 87

Phase three 88

Estimating the Disease Group 88

Estimating the Diseases Individually 90

Discussion 94

Pain Distribution 97

Sensitivity, Specificity, Reliability and Algorithm 100

Pain maps 104

Limitation 109

Conclusion 112

Further Use 113

References 114

Appendices 126

Letters on Ethical Considerations 126

Literature Search Terms 127

Literature Search Results of Databases 127

Shoulder Pain Radiation Form 131

Colour-coded Pain Maps for Common Shoulder Disorders 132

Shoulder Pain Mapping Algorithm and How-to-read the Map 136

Full Statistics reports of the first phase 139

Statistics Reports of the Second Phase 147

Second Phase: Agreement of Estimation and Diagnosis 147

Second Phase: Sensitivity, Specificity 148

Second Phase: Disease Group-Map Group Agreement 154

Second Phase: Group Estimation – Sensitivity, Specificity 155

Statistics Reports of the Third Phase 156

Third Phase: Estimation of the Disease Groups by Testers 156

Third Phase: Disease Groups; Sensitivity, Specificity 160

Third Phase: Group Agreement Between the Testers 166

Third Phase: Estimation of Individual Shoulder Disorders by Testers 169

Third Phase: Sensitivity, Specificity of Estimations of Individual Disorders by the Testers 172

Third Phase: Agreement of Estimations of Individual Shoulder Disorders Between the Testers 193

List of the Tables

i- Table i: Search terms

ii- Table ii: Mean ages for each disorders (phase one)

iii- Table iii: Types and radiation of shoulder pain- phase one

iv- Table iv: The mean of total number of cells marked by the patients and the severity of the pain (VAS) for each disease.

v- Table v: Mean ages for each disorders (phase two)

vi- Table vi: Further imaging in addition to the investigations on the first day

vii- Table vii: Treatments of patients shoulder disorders- phase two

viii- Table viii: Prediction of diagnoses by the researcher

ix- Table ix: Sensitivity and specificity of estimations for each shoulder disorders

x- Table x: Overall shoulder pain distribution (without pain type description) – second phase of the study

xi- Table xi: Types of shoulder pain and regions anteriorly, second phase

xii- Table xii: Types of shoulder pain and regions posteriorly, second phase

xiii- Table xiii: Numbness / pins & needles in the wrist and hand

xiv- Table xiii: Estimation of disease group, sensitivity and specificity

xv- Table xv: Estimation of disease groups by three testers

xvi- Table xvi: Sensitivity and specificity for the disease groups

xvii- Table xvii: Agreement of the raters on the disease groups

xviii- Table xviii: Best estimation of individual shoulder disease by raters

xix- Table xix: Sensitivity and specificity results of testers for individual shoulder disorders

xx- Table xx: General agreement over all (7) categories with 3 raters per disease

xxi- Table xxi: Raters agreement on the estimations of the individual shoulder disorders

List of the figures

I- The sample map to be marked by the patients

II- The Establish Pain Maps for Common Shoulder Disorders

- II-i: ACJ pathology
- II-ii: Instability
- II-iii: Calcific Tendinitis
- II-iv: Rotator cuff pathology (iv-a: phase one, iv-b: phase two)
- II-v : Frozen shoulder
- II-vi: Impingement syndrome
- II-vii: GHJ arthritis

III- Shoulder Pain Mapping Algorithm

Acknowledgement

I would like to acknowledge the support I have received from my supervisors, the college supporter staff and also from the testers, who helped in the third step of the study, as well as the authors of the published part of the study, who contributed to the first study.

Declaration

This thesis has been composed by myself and has not been submitted, in any form, for the award of a higher degree to any other educational institution.

However, the first part of the study has been published in American Journal of Orthoapedics (AJO) (Bayam, Ahmad, Naqui, Chouhan, & Funk, 2011). In addition, the same part was presented at ISAKOS May 2007, Italy, 10th Polythematic Congress of Orthopaedic Surgery and Traumatological Society of Crete in collaboration with British Trauma Society, May 2007, Crete. The paper won first prize as the 'best adult Clinical presentation'- Sir John Charnley Prize at North West Orthopaedic Association Annual Meeting 16th Nov 2007.

Common Abbreviations in the Text:

AC (J): acromioclavicular (joint)

ANOVA: Analysis of variance

CI: Confidence intervals

CT: computed tomography

GHJ: Glenohumeral joint

κ: kappa

MRI : Magnetic resonance imaging

OA:osteoarthritis

ROM: range of movements

R&D: research and development

SPSS: Statistical Package for the Social Sciences

VAS: visual analogue scales

Abstract

Background:

Pain mapping for specific disorders was described in the literature for face, back and hip pain, but not for shoulder pain.

The aim of the study was to fill the gap in assessing patients as a whole for common shoulder disorders, to develop a pathway from subjective experience of patients to diagnoses of pathology, to ascertain specific patterns of pain in patients with common shoulder disorders, to describe comprehensive shoulder pain maps and to test these.

Method:

The study was designed in three phases with prospective blinded method. The first phase aimed to establish the pain patterns for common shoulder disorders. The patients, who presented as new patients with shoulder pain to the outpatient department, were given a custom-made shoulder mapping form to mark their pain, its character and severity. The patients' final diagnoses were coded after investigations and the codes were correlated with the pain map patterns to achieve the aim. Later, colour-coded maps were established for each shoulder disorder. SPSS (statistic package) was used for the first study.

The second phase was designed to test the accuracy of the previously established colour-coded pain patterns in the first phase, to assess sensitivity and specificity of the maps for

the disease groups and each individual disorder, to improve the previous pain mappings and to establish an algorithm. This was achieved by collecting the maps from a larger number of the patients than the first phase's number and the researcher, who was blind to the diagnoses, gave his estimations for each map immediately after collecting the maps from each patient. After all the investigations, treatments and follow-ups, the final diagnoses were coded.

The final phase was to assess inter-tester reliability. The third phase was to test inter-tester reliability of the maps by estimating the shoulder diagnoses using algorithm and colour-coded maps by three raters. This test was used to observe a score of how much consensus or homogeneity there was for the algorithm of the shoulder pain maps. Another aim in the third phase was to examine if the mapping system is easy to use or requires a lot of training.

Statsdirect and VassarStats were used to analyse statistical data in the second and the third phases of the study.

Ethical opinion was sought from the local R&D department and obtained. There was no conflict of interest.

Results:

The first phase of the study included 94 patients and showed that there were definite patterns for each shoulder disorder and it described colour-coded shoulder pain patterns according to the radiation of the pain around and beyond shoulder for six shoulder

disorders; acromioclavicular joint pathology, instability: Bankart's, SLAP etc, calcific tendonitis, rotator cuff pathology, impingement syndrome, gleno-humeral joint arthritis. This showed a range from a very localised pain such as ACJ pathology to a very wide-spread pain such as GHJ arthritis.

The second phase of the study included 194 patients and it tested the mapping patterns from the first phase. The accuracy for the first estimation for individual disorders was 45.4% and the overall accuracy for both estimations was 62.4. The sensitivity was high especially for instability and it was good for ACJ pathology and impingement syndrome. This phase clarified the pain patterns further and detailed three groups of pain patterns. The first group of diseases showed a localized pain around the shoulder and second group showed radiation of the pain beyond shoulder. Later, the second group was subdivided into two. Group 2A showed the pain radiation down to elbow level, whereas group 2B showed radiation below the elbow level.

Third phase was to test inter-tester reliability of the maps by estimating the diagnoses derived from the maps by 3 raters. It tested the reliability for each disease group and individual disease. It showed a substantial agreement between the raters (Kappa (κ) = 0.71)

Conclusion:

A definitive pattern of pain distribution and specific types of pain were demonstrated for common shoulder pathologies. Testing the established maps indicated that the colour-coded maps were reliable and the algorithm was easy to understand. The study advocates

the use of pain maps as an adjunctive diagnostic tool in general practice clinics and orthopaedic / shoulder clinics.

Introduction

Shoulder pain is a common presenting complaint in orthopaedics and reaching a clear diagnosis requires many aspects of a patient's assessment to be explored. There was a gap in the literature for assessing patients with common shoulder disorders as a whole and in developing a pathway from subjective experience of patients to diagnoses of pathology. Therefore, the current study endeavours to establish specific pain patterns for common shoulder disorders as an adjunct to the assessment of shoulder. Conditions causing shoulder pain are common and contribute substantially to the musculoskeletal morbidity of the community (Bjelle, 1989; Urwin et al., 1998). The prevalence of shoulder disorders has been reported to range from 7 to 36% of the general population (Lundberg, 1969) and it is the third most common cause of musculoskeletal consultation in primary care. Approximately 1% of adults consult a general practitioner with new shoulder pain annually (Mitchell, Adebajo, Hay, & Carr, 2005; Urwin et al., 1998).

Pain mapping for specific disorders has been described for face, back and hip pain, but not for shoulder pain. Pain maps were found to be useful, for diagnostic, therapeutic, prognostic and research purposes. Although common pain patterns were described for the common shoulder disorders these are mainly anecdotal in the literature and no study specifically compared and mapped the common shoulder pain pathologies such as subacromial impingement, rotator cuff tears, glenohumeral joint (GHJ) arthritis and acromioclavicular joint (ACJ) pathology. Whilst shoulder diagnosis may involve a complex diagnostic process such as examination, specific tests, baseline x-ray, steroid injection, ultrasound, MRI scan etc., shoulder mapping could be an adjunction to this process.

The initial aim of this study was to ascertain specific patterns of pain in patients with common shoulder disorders and describe a comprehensive shoulder pain map as well as to develop colour coding for these patterns. Secondly, it was aimed to test accuracy of colour coding in the new patients' clinic setting in correlation with the established pain mapping, to improve as needed and develop an algorithm. The objective was to be able to understand if they could have a clinical value in day-to-day practice. The final part aimed to test reliability of the established pain mapping patterns and its algorithm. The objective was used to observe a score of how much consensus or homogeneity there was for the algorithm of the shoulder pain maps.

Nature of Literature Review

Pain mapping for specific disorders was described in the literature for face, back and hip pain, but not for shoulder pain.

The aim of the study was to fill the gap in assessing patients with shoulder problems as a whole, to develop a pathway from subjective experience of patients about their pain, pain types, radiation and severity of the pain to diagnosis of pathology, to ascertain specific patterns of pain in patients with common shoulder disorders, to describe comprehensive shoulder pain maps and to test these.

Aim of the literature search

The aims of the three phases in the current study were different. The literature search was designed to locate the studies for all three phases separately yet, more focused on mapping of shoulder pain in common shoulder disorders as the overall aim of the all three phases was to ascertain specific patterns of pain. Therefore, literature search aim to capture the studies related shoulder pain, diagnosis of the disease and pain mapping for the first phase as detailed below. For the second and third phase, the literature search aimed to review sensitivity and specificity, and reliability of mapping. The literature search for all the phases aimed to locate high quality reviews such as Cochrane reviews and in the absence of those, it aimed to locate any related publication to expand the search including Google Scholar. The aim of the current study was not to perform a systematic review and there were not enough studies in the literature to review on this subject. Nevertheless, the qualities of the relevant studies were checked using critical appraisal questions.

The aim of literature search for the first phase was to locate the studies about general shoulder problems, their diagnoses, investigations leading to diagnoses, and classifications of shoulder problems. By that way, this step aimed to decide if there was any need for pain mapping. In the next step, it aimed to locate previous studies on pain mapping, importance of it and existing literature.

The aim in the second phase was to search the literature about testing the maps and to verify the clinical value and meaning of pain maps once the first phase was completed. In that context, it was to determine the accuracy of the pain mapping and its potential contribution to clinical practice. The literature search also aimed to locate similar studies showing sensitivity-specificity of the subjective experience of patients.

The aim in the third phase was to search how reliable such a tool could be for diagnostic purpose by assessing inter-tester reliability and again to locate similar studies in the literature.

Terms

The main concepts searched in the first phase of the study were:

1. concept: Shoulder
2. concept: Pain
3. concept: Disorder, Disease
4. concept: Mapping

The set of terms were designed to capture all relevant literature related to shoulder pain mapping in the first phase using synonyms and thesaurus from Microsoft Word (Microsoft, 2010) and in addition, mapping and thesaurus, Medical Subject Headings (MeSH) from Medical Literature Analysis and Retrieval System Online (MEDLINE) via National Health Service (NHS) library (ProQuest, 2010;NHS Evidence, 2009). Also, the UK and US English spelling of the words were checked for full terminology. Additionally, this approach was discussed with a local hospital librarian to avoid any major omissions. In the next stage, a table of a list of terms has been prepared to use as a template to search all the databases (table-i in appendices). The other concepts such as sensitivity-specificity of pain maps and inter-tester reliability were searched separately.

Databases Search

In particular, the following databases were searched for relevant studies:

Via the Cochrane library, systematic reviews and protocols have been searched. MEDLINE (from 1950 to 04/December/2013), Excerpta Medica dataBASE (EMBASE) (from 1980 to 04/December/2013) and Cumulative Index to Nursing and Allied Health Literature (CINAHL) (from 1981 to 04/December/2013) database have been searched via NHS library to find the studies of evidence-based medicine, clinical trials, case reports with literature reviews. Also, Google scholar was used and hand search was done in the local hospital libraries (which are not included in the appendix). The searches were not limited by study design or language of publication to maximise the sensitivity of the search. No specific time limit was used. Duplicate results were removed from each individual database search results. The full strategies of shoulder pain mapping are shown in the appendix.

Diagnosis of Common Shoulder Disorders

The population in the current study was the adult patients with shoulder pain, who presented to outpatient clinics as new patients. It can be a great challenge to diagnose the cause of shoulder pain due to the large number of aetiologies. The aetiologies are wide ranged and vary from a minor trauma, which can cause a sprain or a simple muscle strain, to a large tear of one of the shoulder stabilizer muscles. Some shoulder pathologies can lead to chronic pain and limitation in shoulder range of motion such as impingement syndrome, adhesive capsulitis, calcified tendonitis, cervical radiculopathy, glenohumeral osteoarthritis, and biceps tendonitis (Schwarzkopf, Oron, & Loebenberg, 2008).

There are a large number of tools for diagnosis ranging from history and physical examination to a wide range of imaging modalities such as x-ray, ultrasound and magnetic resonance imaging (Schwarzkopf et al., 2008). Yet, owing to the complex anatomy of the shoulder and the spectrum of underlying disorders, the cause of pain can be still difficult to diagnose despite the improvement in imaging technology. Symptoms and medical imaging may not correlate well (Carter et al., 2012).

The mixed or complex shoulder problems might cause difficulty in assessing the shoulder (however, there is no clear literature on how frequently multiple shoulder pathologies occur over the single ones). Dinnes, Loveman, McIntyre and Waugh (2003) stated that there are no clear national guidelines for the diagnosis of shoulder pain. Several diagnostic tests are used for the diagnosis of soft tissue disorders, including clinical assessment, plain x-rays, ultrasonography, magnetic resonance imaging (MRI), magnetic resonance arthrography (MRA) and arthroscopy, yet their relative accuracy, cost-effectiveness and

impact on quality of life are uncertain (Dinnes et al., 2003). The clinical assessment also includes collecting demographic information and assessment of the range of motion, specific physical signs, strength, and stability (Richards et al., 1994).

Although the diagnosis of shoulder disorders should not be based on clinical examination alone, if performed with suggested standardisations, some of the tests are highly reproducible and therefore reliable to use in clinical practice (Johansson & Ivarson, 2009; Guanche & Jones, 2003; Ure, Tiling, Kirchner, & Rixen, 1993). Neer impingement sign, Hawkins-Kennedy impingement test and Jobe supraspinatus test are well-described. The O'Brien's sign is helpful for diagnosing superior labral detachment (Tzannes & Murrell, 2002). The combination of the Hawkins-Kennedy impingement sign, the painful arc sign, and the infraspinatus muscle test yielded the best post-test probability up to 95% for any degree of impingement syndrome (Park, Yokota, Gill, El Rassi, & McFarland, 2005). Streoid injection is used for both diagnostic and treatment purpose in impingement syndrome (Fraser-Moodie, Shortt, & Robinson 2008). The current study included all the physical tests by experienced clinicians in the assessment of shoulder and injection especially for impingement syndrome and ACJ pathology as detailed in the method section.

Shoulder pain associated with rotator cuff disorders and glenohumeral OA can be diagnosed in the majority of patients on the basis of medical history, focused physical examination, and plain film radiographs (Meislin, Sperling, & Stitik, 2005)**.** They are specifically more helpful in diagnosing ACJ pathology and GHJ arthritis. Although plain radiographs can reveal degeneration of AC joint and GHJ arthritis, diagnosis cannot be based on this alone because similar radiographic findings can be seen in asymptomatic

individuals (Mall et al., 2013; Woodward & Best, 2000). In GHJ arthritis, an AP view of the glenohumeral joint will usually reveal degenerative changes and loss of joint space (Woodward & Best, 2000). Similarly, positive radiographs can be also helpful to diagnose calcific tendinitis (Burbank, Stevenson, Czarnecki, & Dorfman, 2008). Antero-posterior, axillary and supraspinatus outlet plain radiography views of the affected shoulder were part of our assessment for each patient.

In Frozen shoulder, radiographs often appear normal as well as CT and MRI, although osteopenia of the humeral head may be noted as a result of disuse (Woodward & Best, 2000). However, arthrography can establish the correct diagnosis of adhesive capsulitis in addition to clinical diagnosis (Brue et al., 2007; R. Neviaser & T. Neviaser, 1987).

The further investigations to diagnose the shoulder disorders include ultrasound, MRI, arthrogram, and computed tomography (CT). Whereas the preferred test for diagnosing rotator cuff disorders is MRI (Dinnes et al., 2003), MRI arthrography has become the preferred test for the imaging of suspected labral pathology (Burbank et al., 2008; Van der Woude & Vanhoenacker, 2007). For full-thickness tears, overall sensitivities and specificities are high with MRI. Where tear prevalence is relatively high, a negative magnetic resonance finding may be sufficient to rule out the presence of a full-thickness tear (Dinnes et al., 2003). Ultrasound was most accurate when used for the detection of full-thickness tears, although sensitivity was lower for detection of partial-thickness tear but specificity remained high (Dinnes et al., 2003). Therefore, it could be more cost-effective in a specialist hospital setting for identification of full-thickness tears (Dinnes et al., 2003). In the current study, ultrasound, CT arthrogram, MRI and arthroscopic

procedures were part of decision process in diagnosing the shoulder disorders as explained in the method section.

The common shoulder disorders included in the current study are impingement syndrome, adhesive capsulitis (frozen shoulder), calcified tendonitis, ACJ pathology, glenohumeral osteoarthritis, rotator cuff pathologies, and instability problems. In addition to their classification, the descriptions and the types of pain they can cause were discussed as below.

Classification of Common Shoulder Disorders and Their Descriptions

There are a large variety of shoulder disorders and some classification systems to categorise these into different groups. Most shoulder problems fall into three major categories: soft tissue disorders, articular injury or instability, and arthritis (Dinnes, Loveman, McIntyre, & Waugh, 2003). One of the studies classified these into six diagnostic categories: capsular syndrome (adhesive capsulitis, arthrosis, frozen shoulder, etc.), acute bursitis, acromioclavicular syndrome, subacromial syndrome (chronic bursitis, tendinitis, rotator cuff tears), rest group, and mixed clinical picture (De Winter et al., 1999), whilst another study described three different patterns of shoulder pain. These were Pattern 1: impingement pain, Pattern 2: acromioclavicular joint pain and Pattern 3: shoulder pain (frozen shoulder; glenohumeral osteoarthritis; complete cuff tear; subscapularis tear; painful laxity; post-traumatic instability; and internal derangement) (Carter et al., 2012). They all have some specific features as well as common characteristics. It is not always easy to diagnose these as they may not be an isolated problem but may be a rather a complex and mixed one.

Subacromial impingement syndrome is a symptomatic diagnosis that may be the result of several patho-anatomical processes. Subacromial impingement syndrome can be described as the compression of the suprahumeral structures against the anteroinferior aspect of the acromion and coracoacromial ligament (Calis et al., 2000). It is suggested there is a multi-factorial aetiology for this disorder. It includes inflammation of the bursa, degeneration or overuse of the rotator cuff tendons, weak or dysfunctional rotator cuff and/or scapula musculature, posterior capsular tightness, postural dysfunctions of the spinal column and bony or soft tissue anomalies (Limb & Hay, 2007).

Glenohumeral stability depends on several factors. These are competent capsulolabral structures, effective muscular activity, intact and effective neural connection and an absence of excessive extrinsic deforming force (Limb & Hay, 2007). Deficiency of these will lead to instability. There is an association between instability and dislocation. Anterior shoulder dislocations typically occur when the arm is abducted and in external rotation. Because the shoulder has an extensive range of movements (ROM), it is at risk for developing instability and is the most commonly dislocated joint in the body (Milewski, Hart, & Miller, 2012). Dislocation of the glenohumeral joint is the most common large joint dislocation with an incidence of 1-2 % (Limb & Hay, 2007). A common reason of laxity is the labral injury. The SLAP lesion (*s*uperior *l*abrum *a*nterior and *p*osterior) and other glenoid labral tears are common in throwing athletes who present with a painful shoulder that clicks or pops with motion (Woodward & Best, 2000).

Calcific tendinitis of the shoulder is defined as a process involving calcium deposition commonly in the rotator cuff tendons. The disease is often chronic in nature and a cell-mediated process. However it is usually self limiting in terms of its acute pain states (Hurt & Baker, 2003).

Milewski et al. (2012) described that rotator cuff disease is a continuum beginning with mild impingement and progressing towards partial tear, full-thickness tear, and finally arthropathy of the rotator cuff. In younger patients, the rotator cuff injuries typically result from trauma whereas in the patients older than 40, chronic impingement syndrome often results in cuff tear (Woodward & Best, 2000). Diseases of the rotator cuff cover a

spectrum of disorders including tendinosis, tendonitis, tendinopathy and tears. Rotator cuff arthropathy is defined as a massive tear of the cuff. (Limb & Hay, 2007).

Another common shoulder disorder is adhesive capsulitis, or frozen shoulder, which is a condition characterised by global restriction in the range of the glenohumeral joint (Limb & Hay, 2007). This pathology is defined as a self-limiting condition of unknown aetiology (Brue et al., 2007). An autoimmune cause has been proposed (Woodward & Best, 2000). The diagnosis is mainly clinical and no significant changes are normally present at MRI or CT scan (Brue et al., 2007).

Glenohumeral arthritis may develop following previous trauma, rotator cuff tear or from underlying causes such as rheumatoid arthritis, Lyme disease. Multiple joint involvement is suggestive of rheumatoid arthritis. Arthritis of the glenohumeral joint generally causes pain with activity, loss of passive motion and stiffness. Some patients may complain of night-time pain (Woodward & Best, 2000).

AC joint problems can be divided two broad groups; AC joint disruption and osteoarthritis (OA) and/or osteolysis of the AC joint (Rull & Colin, 2013). Whilst AC joint disruption are seen very often as athletic injuries as a consequence of fall, direct blow or repetitive overhead motions, OA may occur after a trauma and more rarely as primary phenomenon. Distal clavicular osteolysis may occur spontaneously in rheumatoid arthritis, hyperparathyroidism, myeloma, systemic sclerosis, due to infection and in those who are involved in throwing sports/extensive upper limb weight training (Rull & Colin, 2013).

Osteoarthritis of the acromioclavicular (AC) joint is a common condition causing anterior or superior shoulder pain, especially with overhead and cross-body activities (Mall et al., 2013). This most commonly occurs in middle-aged individuals because of degeneration to the fibrocartilaginous disk that cushions the articulations. Diagnosis relies on history, physical examination, imaging, and diagnostic local anesthetic injection (Mall et al., 2013). The patients with all the above disorders present to the clinics with shoulder pain, which requires further assessment.

Importance of Mapping to inform diagnosis

The pain was described as an experience that was perceived by the patient (Pain management, 2014). Another definition was “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage" by International Association for the Study of Pain (IASP) (International Association, 2012). Pain is the most common symptom in musculoskeletal medicine and is described using a variety of terms. Clear differences exist between the “throbbing” pain of an abscess, “aching” pain of chronic arthritis, “burning pain” of neuralgia and the “stabbing pain” of a ruptured tendon. The precise location of pain is important in orthopaedics but does not always correlate with the site of pathology (Rodowsky & Bigliani, 1997). In the spine, pain signals are generated from nerve roots and travel neural pathways through the body. When nerve roots are compressed, they generate pain signals that can be felt in different areas throughout the body (North American Spine, 2013). However, pain arising in or near the skin is usually localised accurately, as is pain from intrinsic shoulder pathology (Rodowsky & Bigliani, 1997) whilst pain arising in deeper structures is more diffuse and sometimes has an unexpected distribution (Apley & Solomon, 1993). One explanation for this in the shoulder is due to its proximal location in the “sclerotome” and the extensive convergence of afferent signals from this region to the dorsal horn of the spinal cord. The “sclerotome” is defined as pain arising within the periosteum and muscle innervated by one spinal segment (Inman & Saunders, 1944).

.

In the literature, there were limited studies, which attempted to specify the pain patterns for shoulder disorders. One of the studies explained that pain patterns can be broadly distributed to the deltoid, trapezius, and posterior scapular region. The location of symptoms may or may not correspond to the proximity of the pain generator (Sizer-Phillip, Phelps, & Gilbert, 2003). Cervical disc disease commonly presents with pain referred to the shoulder. This pain is most often referred to the posterior aspect of shoulder and trapezius and occasionally to forearm or hand (Rodowsky & Bigliani, 1997). Other examples of referred pain involving the shoulder include the stimulation of the diaphragmatic tendon centre that typically produces shoulder pain (Di Massa, Avella, & Gentili, 1996).

Similarly, shoulder pain with radiation to the arm and hand in an ulnar nerve distribution could be an indicator to the existence of a Pancoast tumour (Kovach & Huslig, 1984). Pain from the sternoclavicular joint can be referred pain to areas distant from the joint (Hassett & Barnsley, 2001).

Many different shoulder disorders cause similar symptoms and patterns of pain (Larson, O`Connor, & Nirschl, 1996). The pain related to rotator cuff pathology is described as insidious onset exacerbated by overhead activities and it is in the deltoid region (Rodowsky & Bigliani, 1997). Arthritis of the glenohumeral joint generally causes pain with activity and frozen shoulder classically consists of shoulder pain that is slow in onset (Woodward & Best, 2000).

Gerber, Galantay and Hersche (1998) aimed to ascertain the distribution of ACJ and subacromial impingement pain by injecting hypertonic saline into the ACJ and

subacromial spaces of normal subjects. There was a description of pain distribution for calcific tendinitis from the point of the shoulder to the deltoid insertion commonly and, less frequently, to the neck (Woodward, 2013),

Subacromial irritation in Gerber's paper resulted in an "intense" pain mainly in the lateral border of the acromion and the lateral portion of the deltoid muscle (Gerber et al., 1998). Dutton (2008) described that the pain that radiates beyond the elbow is far less likely to be due to shoulder pathology, particularly if it is associated with any sensory disturbance in the limb such as distal radiation or pain, numbness or paresthesia (Dutton, 2008). Similarly, Woodward and Best (2000) concluded that the pain related to impingement usually occurs over the anterolateral aspect of the shoulder, often with some radiation to, but not usually beyond, the elbow.

The patient's experience of the pain is the key many times. It is widely taught that diagnosis is revealed in the patient's history and even the history alone may display the diagnosis. Sometimes it is all that is required to make the diagnosis (Rull & Draper, 2011). Pain mapping is illustration of patients' statement of their subjective complaints and possibly a part of history taking. Therefore, a well-designed visual pain map may guide patients to express their symptoms in a more descriptive way and in return, this may help doctors for diagnoses.

Existing Literature in Pain Mapping

Modern pain mapping was introduced in 1949, when Palmer (1949) provided outline diagrams of the human body and asked patients to mark on the charts wherever they experienced pain. The usage of pain maps in clinical practice is now more widespread and forms part of the McGill pain questionnaire (Rankine, Fortune, Hutchinson, Hughes, & Main, 1998). When Rankie et al (1998) assessed the ability of the pain drawing to predict the presence of nerve root compression, there was considerable overlap in the appearances of the pain drawings between patients with and without nerve compression. They concluded that pain drawing should be interpreted with caution and in light of the full clinical picture. Accordingly, our approach to pain mapping was to examine if the pain mapping can be used as adjunct to the full assessment in shoulder disorders' diagnoses on oppose to using as sole diagnostic tool.

Patients can complete the self-evaluation portion of the questionnaire in the absence of a physician. A shoulder score may involve the visual analogue scale score for pain and the cumulative activities of daily living score (Richards et al., 1994). The questionnaire approach may give a fairly good picture of the neck/upper extremity status (Ohlsson, Attewell, Johnsson, Ahlm, & Skerfving, 1994). The primary types of self-reported measures of pain include the verbal rating scales (VRS), numerical rating scales (NRS), visual analogue scales (VAS) and pain drawings (PD). Each of these methods has advantages and disadvantages (Disorbio & Bruns, 1999). Despite the simplicity, a number of persons had difficulty using them however; the VAS and NRS scales are more easily comparable. The article showed different aspect of the pain scales and suggested that there

may be an identifiable pattern of specific pain ratings that are consistent for each medical diagnosis. The paper's review on VAS and suggestions were relevant to our approach of specific pain patterns for each shoulder disorder as well as the choice of additional use of VAS in our custom-made shoulder mapping form.

The pain mapping has been described for neck, face, back and hip pain. Different projects of using pain maps are suggested. One of them is to provide markers for the location of pain in the human brain (Medicalxpress, 2012). It is also suggested that pain map could be used to determine the kinds of pain which are felt by patients who are unable to articulate it. This could be facilitated using functional magnetic resonance imaging (fMRI). The study named the pattern, which is obtained following a charted blood flows through the brain, as the signature of brain activity (Pain Signature, 2013). Another project is three-dimensional pain mapping using a computer method, which allows patients to mark the location, intensity, and depth of pain on a three-dimensional model of the human body (Baird, 2009). The model can be rotated to obtain better viewpoints for marking pain.

No other study was found in the literature to test shoulder pain mapping as the current one was the first study. Among the limited literature, one study was by Pang et al. (1998) on analysis of spinal pain mapping. The authors tried to find the source of pain and map the pain related to back pain after all the investigations. The study was not a blinded study with potential bias. The level of the evidence for this study was four. It concluded that spinal pain mapping is a functional approach to the diagnosis. However, it was performed in conjunction with nerve block and it was to analyse the source of pain including the other investigations including CT, MRI and EMG. The study by Turp et al. (1998) was

about pain maps from patient with persistent facial pain. The aim was to look the facial pain radiation and how the pain distribution was related to the dermatomes. The study was a prospective and primary cohort study. The design was appropriate for diagnostic purpose but there was not any explanation about ethical implications. The level of evidence was 4 for this research.

The second and third phases of the study were related to sensitivity and specificity of a clinical measurement. Sensitivity shows the proportion of the patients with the disease, whose tests are positive and specificity shows the probability that the patients does not have the disease when they are disease free (Faragher, 2005). In the literature, there were no other studies on the sensitivity and specificity of patients' subjective assessments in diagnosing the shoulder diseases.

The aim of the third phase in the current study was to test inter-tester reliability which helps to understand the consistency of the raters estimates (Cherry, 2013). This test was used to observe a score of how much consensus or homogeneity of algorithm of the shoulder pain mapping. In the literature search, most of the relevant intertester reliability studies were on spinal pain and there was one study on shoulder pain. The aim of the study by Carter et al.(2010) was to describe and determine the inter-tester reliability of a newly developed classification system of shoulder syndrome recognition with inter-tester reliability on 255 patients. It was designed as a blinded study and the kappa coefficient was 0.664. One of the spine studies had an aim to determine the inter-tester reliability of a low back pain classification system among experienced and novice clinicians. This was a prospective cohort study and included 204 patients. The kappa coefficient was 0.61. Those

studies help us in designing the third phase of the current study however, those two studies were not related directly to pain mapping.

Literature search did not locate another study on shoulder pain mapping for common shoulder disorders and it did not reveal any other study for sensitivity-specificity of pain mapping and inter-tester reliability.

Identification of Necessity for this study

The literature search showed no previous study on pain mapping for shoulder disorders. There was a gap in the literature in assessing patients with shoulder problems as a whole from the patients' subjective description to diagnosis of pathology in their shoulder.

The experience and observation of the authors in the first phase of the study led to believe that there is a definite pain pattern for the shoulder disorders and the literature showed that pain mapping as a tool is good way of assessing the pain and VAS is a useful tool for the pain severity. (Disorbio & Bruns, 1999 ; Rankine et al., 1998). Therefore, custom-made shoulder mapping was designed for the current study (figure I).

This study potentially would help us to understand the pattern recognition in shoulder pain and direct the examiner for use of specific test to support the diagnostic hypothesis generated by maps / patterns. Although it would not be enough to establish the accurate diagnoses in the absence of other clinical assessment tools but may strongly suggest the one. This could be similar to the example that O'brien test is used in diagnosis of both SLAP lesion of shoulder instability and ACJ pathology but more pathognomonic for ACJ pathology. Similarly, the shoulder pain mapping may show a pattern which may fit to two different diagnosis but could fit one diagnosis more than another one.

This study aimed to fill this gap with a visual questionnaire, which was easy to understand and mark. This was probably to aid the diagnosis as an adjunct to the other assessments. Should the study show the definite patterns, testing these would help strengthen the aims

of the study. Therefore, the second and third phase of the study was planned to test the established shoulder pain patterns via blind estimation of the diagnoses and inter-tester reliability by researcher and other raters.

In summary, the aim of the current study was to ascertain specific patterns of pain in patients with common shoulder disorders, describe a comprehensive shoulder pain map and to test their reliability.

Method

Pain mapping for specific disorders was described in the literature for face, back and hip pain, but not for shoulder pain. Literature search revealed a gap on this area. The aim of the study was to fill the gap in assessing patients as a whole for common shoulder disorders, to develop a pathway from a subjective experience of patients to diagnosis of the pathology, to ascertain specific patterns of pain in patients with common shoulder disorders, to describe comprehensive shoulder pain maps and to test these. This aim of the method section was to detect which methodology and approach to be used in order to achieve the aims.

It was planned as a prospective blinded study, which was conducted in three phases. Each phase had some differences in terms of their methodological approach. Therefore, this section was divided into three parts and each phase of the study was explicated separately in details. Each part included the following headings: method, design, participants, recruitment, procedure, data analysis, ethical implications.

Phase One

The first phase was to establish pain patterns for common shoulder disorders. The aim was achieved by collecting custom-made pain maps (which was explained in details as below) from the patients, who presented to the shoulder outpatient clinics as new patients and by comparing them with their final diagnoses. The details of the first phase were explained in method, design, participants, recruitment, procedure, data analysis, ethical implications sections.

Method:

The method for the first phase of the study was planned as prospective and blinded one. A prospective study usually involves taking a cohort of subjects and watching them over a long period. Comparing to retrospective studies, however, prospective ones generally have fewer potential sources of bias and confounding factors. Most sources of error are due to confounding factors and bias is more common in retrospective studies than in prospective studies (StatsDirect, 2013).

.

Blinding is another key factor in studies in terms of their bias. Hróbjartsson et al. (2013) explained that non-blinded outcome assessments were very common in orthopedic traumatology and they tended to generate substantially biased effect. In some trials, conscientious non-blinded assessors may overcompensate for an expected bias in favour of the experimental intervention and paradoxically induce a bias favouring the control, whereas other trials would have fairly neutral assessors with no important bias (Hróbjartsson et al., 2013). Blinding of the researcher to the diagnoses in the current study might have an important role to prevent the bias about their estimations.

The study aimed to perform a quantitative study about shoulder pain mapping. However, the pain can be subjective and the reception of the pain may vary from patient to patient. An addition of qualitative assessment to the study about pain perception of the patients would increase the strength of the study and potentially give a better picture of patients' experience and pain patterns. However, this will probably mean further extension of the study and this could be beyond the scope the current study.

Design:

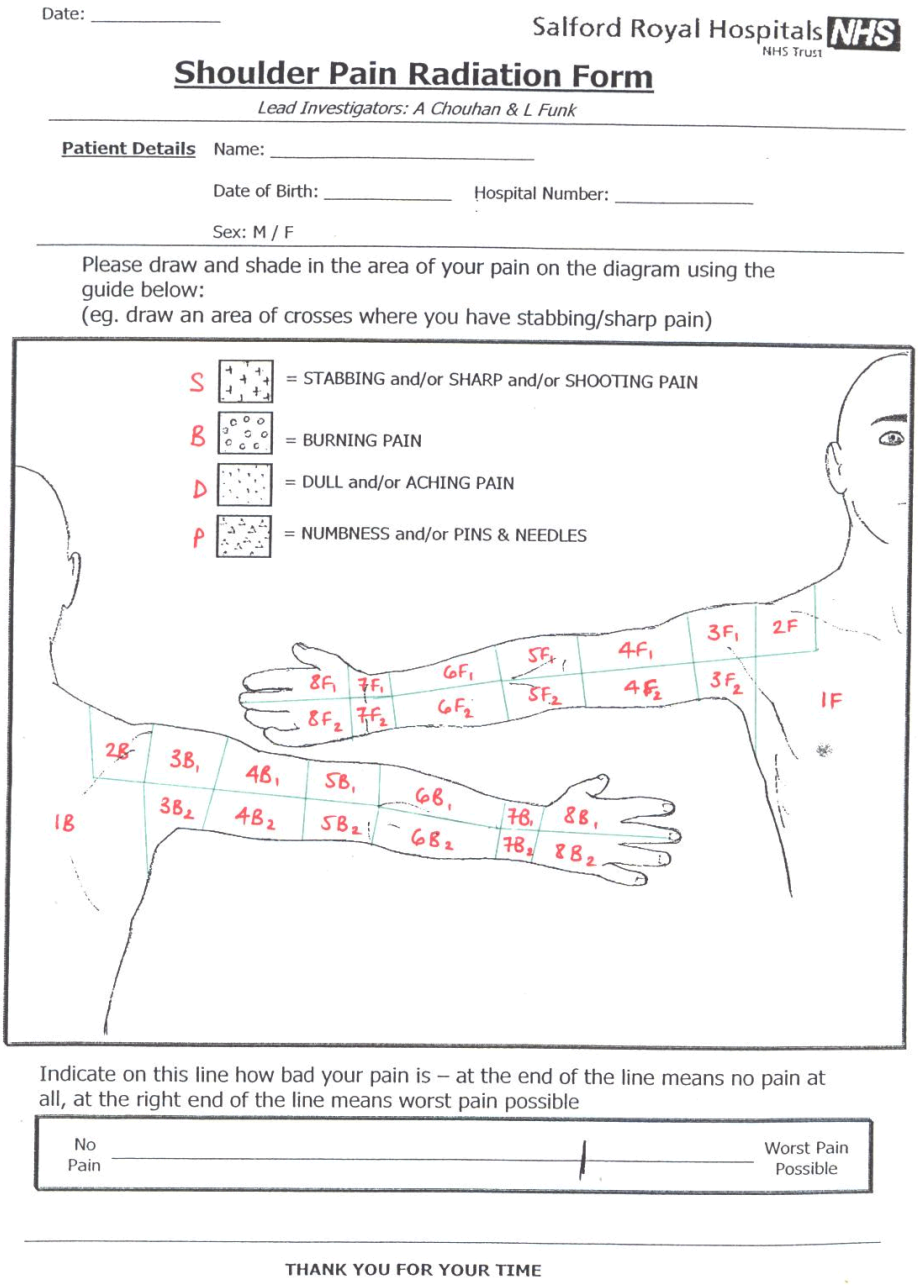
This was the first phase of the three-phase study. The phase one aimed to ascertain definite pain patterns of common shoulder disorders via a shoulder pain radiation form, which was a custom-made pain map (as explained below) showing pain localisation, type of pain and severity and collected from each patient (Figure I). Then, the maps were correlated with the final diagnoses which were coded after all the investigations.

The custom-made shoulder map was developed to simplify the appearance of the shoulder with the pain types and also visual analogue scale (VAS) was added for the severity of the pain so that the patients would understand without detailed explanation in the outpatient setting (figure I). in the end, custom-made upper limb pain map was a shoulder pain radiation form, which was used in the study illustrated the type of pain, its severity and area/s of radiation. The pain map showed the anterior and posterior part of the whole arm including the neck and shoulder. Each side of the arm was divided into 14 sections or cells giving a total of 28 cells. The acromio-clavicular joint and axilla regions were included amongst the cells. A choice of 4 different pain types, each with an associated symbol were given to the patients to illustrate their pain on the pain map; + indicated a sharp, stabbing

or shooting pain, o for burning pain, • (dot) indicated a dull or aching pain and Δ for numbness and pins & needles. For analysis, the abbreviations; S, B, D, P were used respectively (figure I).

Severity of the pain was assessed using a visual analogue scale (VAS), which was added to the custom-made upper limb pain map (figure I). Patients rated the intensity of pain on a continuum of no pain to the maximal worst pain imaginable. The VAS score is the distance from the lowest pain level to the mark made by the patient (Jensen & Karoly, 1991).

Several studies used pain mapping and VAS questionnaire. For example Rankine et al (1998) used pain drawing in a prospective study and they compared the pain drawings results with MR lumbar spine magnetic resonance imaging. In the current study, we conducted a prospective study similarly, and we compared the pain mapping with final diagnosis including Ultrasound, x-rays, MR,and clinical assessment. We used multiple tools for final diagnoses as some shoulder disorders can be difficult to diagnose and only MR scan would not be enough for full diagnoses with false negative rates (Dinnes et al., 2003).



Date: _____

Salford Royal Hospitals NHS
NHS Trust

Shoulder Pain Radiation Form

Lead Investigators: A Chouhan & L Funk

Patient Details Name: _____

Date of Birth: _____ Hospital Number: _____

Sex: M / F

Please draw and shade in the area of your pain on the diagram using the guide below:
(eg. draw an area of crosses where you have stabbing/sharp pain)

Indicate on this line how bad your pain is – at the end of the line means no pain at all, at the right end of the line means worst pain possible

No Pain ——— Worst Pain Possible

THANK YOU FOR YOUR TIME

Figure I: The sample map to be marked by the patients (excluding the red and green marks: they were only used on comparison sheet)

Participants:

The participants were to be adults (16 years and over) with shoulder pain and the inclusion and exclusion criteria as below.

Inclusion Criteria: The patients, who were adult (16 years old and over), presented to the shoulder clinic of a university hospital as new patients with shoulder pain but without any previous intervention or previous treatment were included to the study.

The patients in the current study were selected from the new patients because if any previous treatment was undertaken for these patients, their pain patterns could be changed due to the treatment.

We did not aim an age limit as different shoulder disorders may present in different ages. The shoulder clinic of the hospital in its nature was designed to see all the adults with shoulder problems. It did not include the population under the age of 16.

Exclusion Criteria: The patients with symptoms of neck pain, clinical features indicating neck pathology such as disc disease, spinal cord compression, or previous and multiple shoulder problems, ipsilateral other upper limb problems were excluded from the study. Additionally, any patients with suspicion of carpal / cubital tunnel syndrome or nerve compression were also excluded from the study as some shoulder pain may present with numbness or pins and needles. After initial exclusion, if the patients had no complete diagnosis, or diagnosed later with neck, multiple shoulder problems and carpal / cubital tunnel and if the patients marked the shoulder maps inappropriately (such as wrong

marking signs, outside the area or no marking at all etc.), they were excluded as well at a later stage, as shown in results section.

Recruitment:

The patients, who presented as new patient with shoulder pain, were recruited from the outpatient clinical of a university (teaching) hospital. The patients gave their verbal consent to mark the localisation their pain, type of the pain and the severity on the map. The recruitment was specifically from the new patients to avoid any mixed picture of the shoulder problem as the previous treatment could mask the pain's character.

Procedure:

All the patients were given a full explanation as to the nature of the study and they gave informed consent to participate at this stage. They were asked to demonstrate and mark their pain type, localisation and severity on the custom-made shoulder pain map. All the maps were coded with a final diagnosis after all the investigations as detailed below. The pain maps with multiple diagnoses or any doubt in diagnosis were excluded.

The specialised clinicians' opinions, who were either upper limb orthopedic surgeon or senior upper limb physiotherapist and who were blinded to the map, were sought for diagnosis. In conjunction with the other investigations, the definite diagnoses were coded to the questionnaire form once all those assessments were completed for each individual patient. It was a definitive end point diagnosis after combination of all the above procedure. The first phase part of the current study was to establish a specific pattern of pain for each common shoulder disorder.

All patients completed an upper limb pain map prior to their consultation with the clinician. Patients were given clear verbal explanations on the map about how to complete it. The ones who agreed to complete the map were included to the study.

In outpatient clinical setting, all the patients were assessed with a detailed history, examination and specific shoulder tests (Apprehension, Hawkin's, Gerber etc.). Radiographic assessments and ultrasound scans were performed as part of the setting in the outpatient clinic. Whilst radiographs included antero-posterior, axillary and supraspinatus outlet views of the affected shoulder, ultrasound scans were carried out vastly on the patients with suspicion of rotator cuff tears, impingement and calcific tendinitis. Some patients had MRI scan, MR arthrogram and / or arthroscopy at a later date before confirming their diagnosis. Local diagnostic injections were also performed in a number of cases especially when clinical diagnosis suggested ACJ pathology, impingement and calcific tendinitis. The figures about the injection and further investigation were added to the result section. There were 94 patients included to the first phase of the study.

Data collection

There were six groups included to the study in the first phase. Collecting the data clinically to get enough numbers for a clinical study is always a difficult task and time taking process especially when the clinicians change their work place very often. Therefore, the number of patients for some groups was limited.

The patient data included all the assessments of the patients from the first presentation to the treatment in the first phase. Data can be collected through the application of techniques such as interviews, questionnaires, observation, direct physical measurement, and the use of standardised tests (Polgar & Thomas, 1991). The role of chance can be minimised by designing a study with adequate sample size and precision to assure a low Type 1 error rate (Newman, Browner, & Hulley, 2001). In the current study, the mapping forms were collected through pain questionnaire forms during their clinical assessment in outpatient clinical setting. Additionally, the data about their diagnosis were obtained through the electronic patient data from the clinical letters, investigations and operation notes. The next step was to code the diagnoses to the maps.

The examining clinicians were blinded to the pain map result. In conjunction with the other investigations, two separate clinicians' opinions were sought for diagnosis in order to help increasing the precision of diagnoses. The definite diagnoses were not coded to the questionnaire form by the researcher until all those assessments completed for each individual patient.

Data analysis (and analytic strategy):

The final diagnoses were correlated with the results of the pain maps. Cells on the shoulder pain radiation form were considered positive when the appropriate symbols had been marked. The distribution of the pain was manually correlated with the clinical diagnosis to ascertain a relationship between site of pain and the particular shoulder pathology.

The data from the custom-made upper limb pain map and the additional data of clinical assessment about the diagnoses were collected and converted to a data sheet (excel) for analysis. Then, the data sheets were transferred to statistical packages as required.

The map patterns were established by using the excel sheets for each diagnosis. Each side of the arm was divided into 14 sections or cells giving a total of 28 cells. On excel, each cell was assessed for each individual diagnosis. The number of the cells, concentration of each pain type in the cell and localisation were assessed as a group map for each diagnosis and compared with the other diagnoses by means of excel sheets and summary tables. Then the pattern for each disorder was established and the differences were evaluated.

One part of the process was to establish colour-coded pain patterns, which meant to facilitate easier understanding of different pain patterns for each shoulder disorder. During the process of submission of the first phase for publication, one reviewer suggested to code pain patterns and to reflect each type of the pain with different colours so that the complex pain patterns on the tables would be distinguished better and would add simplicity to recognition. Therefore, a colour-coded shoulder pain patterns was established for each shoulder diagnosis and this included the pain type and distribution on the upper limb. The final patterns representing each disorder of the upper limb were mapped anteriorly and posteriorly.

However, there were other possible options of establishing these maps. One of them could be a statistical approach to correlate the diagnoses with each cell, pain type or distribution. Nevertheless, due to high number of the cells and the pain types as well as the varieties in

patients' descriptions of the pain on the map, it looked very complex and needed more patients in each group. In addition, the opinion from a local statistician supported this view. Therefore, statistics was not used to establish pain mapping patterns. Yet, the number of cells marked (distribution of the pain) by the patients and severity of the pain were able to be tested statistically.

The number of the groups and the number of the patients in each group were not known initially and this was not predicted or calculated prior to the study as the researcher was blind to diagnoses. The final diagnoses were obtained at the end of phase one and the number of the groups and the number of the patients in each group were determined at this stage. Therefore, sample size was not calculated prior to the study. But post-hoc calculations were performed for the results with statistically significant findings.

As the number of the area, severity and type of pain were quantified, the study type was a quantitative study and the data was parametric with more than two groups in the study. Statistical analysis was conducted using SPSS (Statistical Package for the Social Sciences) 16.0 version for Windows included P-P plot to determine if the sample is normally distributed and Pearson's correlation tests to see any correlation between the number of the area and the severity of pain in the first step of the study. On assumption of normal distribution, one-way ANOVA (the analysis of variance) was used.

The decision on choosing the test (ANOVA) is based on one independent variable and multiple dependent variables. It is not appropriate to treat each pair-wise testing as a single test when there are multiple of them, using the conventional 5% level of the significance (Anthony, 1999). The solution of revised α value for multiple comparison is to use

ANOVA for parametric data instead of *t*-test. The important items to note are the *F* value and its significance. The *F* value increases as the groups differ more (Anthony, 1999).

However, the rest of the data such as pain patterns, which was not comparable statistically due to high number of varieties (several diagnoses with several types of pain) and comparatively low number of patients, was analysed over excel sheet as "hand comparison" and the general patterns of pain character and distributions for common shoulder disorders were established.

One-way ANOVA was used to find a meaningful difference in the number of areas marked by the patients. The VAS scores were analysed to see if there is any significant difference in the range of pain between the diagnoses. Opinion was sought from the local statistician for statistical analysis.

Upon completion of first phase of the study, the colour-coded shoulder pain maps for common shoulder disorders were achieved.

Ethical implications

The current study was conducted in the outpatient clinic setting by providing a map to the patients to mark their pain type, distribution and the severity. The patients presented to the hospital with a new shoulder problem and there was not any intervention by means of an additional investigation or a treatment for the purpose of the current study. National Research Ethics Service (NRES) advises to seek opinion from research and development

(R&D) office clinical governance office in the first instance (Health Research, 2013). Ethical opinion was sought from local R&D department. The local R&D department gave their opinion saying that ethical approval is not needed for this study as all the investigations and the treatments included in the study were already part of their routine clinical assessment, and there was not any additional investigation or intervention related to the current study.

The other areas, which might cause potential problems or dilemma, are anonymity of data, public availability, and possible future prospective studies, which the current study could lead to. Regular revision may be required to avoid breaching ethical limits at any stage.

The data of the current study was anonymised by giving a number to each patient and matching identifications of patients in a different file, which was held in a separate document folder. To further enhance security, the memory device, which was used for the anonymised data of the study, was encrypted.

Conflict of interest is another ethical issue for a study. Participants in peer review and publication should disclose their conflicting interests, and the information should be made available, so others can judge their effects for themselves (International Committee, 1993). There was no conflict of interest for the current study and participants had no financial interest in any product or service related to phase one the study.

In summary- Phase one

Population: Adult patients (16 years and over), presented as a new patient with shoulder pain.

Intervention: There was no intervention related to the study. The investigations in outpatient clinic were part of routine patient assessments

Comparison: Pain patterns and common shoulder disorders

Outcome measures: To ascertain specific patterns of pain in patients with common shoulder disorders, to establish an algorithm to facilitate the use of maps in practice

Phase Two

The aim of the second phase to test the color-coded pain patterns which were established during the first phase of study, to analyse their accuracy, sensitivity-specificity, to improve them if required and to establish an algorithm. The clinical setting was the same as the first phase but hospitals and the sequence of the procedures were different. The aim of this phase was achieved with another prospective study with a larger group of new patients, who marked the localisation of their pain, type of their pain and severity on the same custom-made shoulder pain map and the researcher estimated the diagnoses of shoulder problem only using the established colour-coded pain maps from the first phase of the study.

Methods:

This was a prospective study as it was done in the same manner as the first phase but with a new group of patients. The researcher was blinded to the diagnoses when he estimated the diagnoses of the shoulder disorder. The researcher coded the provisional diagnoses immediately after each patient handed their pain maps.

Design:

In the second phase of the study, as in the first phase, the patients, who presented to the shoulder outpatient clinics, marked their pain localisation, type and severity on the same custom-made maps. Different to the first phase, the second phase aimed a larger group of patients to increase sample size for each disease group, for better analysis and statistical evaluation and the researcher gave two estimations of diagnoses for each map immediately after collecting from the patients. The researcher used only previously established colour-

coded maps for the estimation without any further information about patients' history, clinical assessment and treatment.

The final diagnoses of the shoulder disorders were coded as explained in the below procedure section, in similar way to the first phase and the estimation and diagnoses were correlated to detect the accuracy of the estimation, sensitivity-specificity

Participants:

The participants were adult (over 16 years old) patients with shoulder pain included to the study with the same inclusion and exclusion criteria; the same as it was in the first phase of the study.

Recruitment:

The patients, who presented as new patients with shoulder pain, were recruited from shoulder outpatient clinics of two different hospitals; a large and sub-specialised orthopaedic hospital and a private hospital. The patients, who were involved in the study, were given a map showing the front and the back of the shoulder and the pain types (figure I). They were asked to mark the pain type, distribution and the severity of the pain before their clinical assessment. 194 patients were included to the second phase of the study.

Procedure:

The patients were asked to demonstrate and mark their pain and severity on the map and they gave informed consent to participate as it was in the first phase of the study. The collection of the maps in the second study lasted 10 months (January 2010-October 2010). As soon as the shoulder pain radiation maps were collected from the patients, two estimations of diagnoses were coded by the researcher, who was blind to the diagnoses, and used the previously established colour-coded shoulder pain maps in the first phase in order to test their accuracy. They were coded immediately after they were marked by the patients and before any further data was obtained about the patients. The reason of coding the maps with two estimations was some similarities in the pain patterns of different shoulder disorders. The shoulder maps were anonymized, filed and kept separately with numbers from 1 to 194.

The second phase of the study was planned to leave longer period before collecting data about the final diagnosis of each patient. By this way, more accurate diagnoses would be able to be obtained. At the end of the long periods (at least12 months, November 2011 to start data collection), all patients had their investigations completed as well as their treatment including operations and follow-ups. Besides, the other aim of this phase was to improve previous colour-coded diagnostic figures and to contribute to the results from the first phase. The data about their diagnosis were obtained through the electronic patient data from the clinical letters, investigations and operation notes.

Data analysis:

Once the final diagnoses were obtained, these were coded to the maps. The diagnoses' codes were correlated with the estimation codes to test the accuracy of the established pain maps for each diagnosis. Then, the two phases were combined to conclude the final patterns for each shoulder disorder.

The excel sheet is transferred to the package for analysis and reports as it was in the first phase. The second study was analysed using VassarStats (Lowry, 2013) and Statsdirect(2013) . and again. Initially, sensitivity and specificity (of 2 groups and) for each diagnoses were checked with VassarStats. The aim of the second phase was to test the previously established pain mapping and to find out how well the estimations agree with the diagnoses. Sensitivity, specificity, and predictive values are used to give the value of the clinical decision tool (Harris & Taylor, 2008) (please refer to statistical terminology). Confidence interval was used to exclude the trivial estimation. Sensitivity and specificity of the disease groups and also for each individual disorder were described. For this purpose, VassarStats was used (Lowry, 2013) and StatsDirect (2013) package was used for analysis of the other data such as age, agreement between estimation and diagnoses, agreement between map group and disease group. Opinion was sought from the local statistician for statistical analysis as it was done in the phase of the study.

An algorithm was aimed to produce at the end of this phase in addition to the improvement in the patterns of pain maps. Furthermore, a how-to-read the shoulder maps guidelines was planned to add to conduct the third phase of the study, inter-tester reliability.

Ethical Implications:

The setting of the study and participants were as in the first study, only the hospitals were different. Ethical opinion was sought from local R&D department and obtained.

The data of the second phase was anonymised by giving a number to each patient and matching identifications of patients in a different file, which was held in a separate document folder. The memory device, which was used for the anonymised data of the study, was encrypted

There was no conflict of interest for the second phase and participants had no financial interest in any product or service related to phase two the study.

In summary- Phase two

Population: Adult patients (16 years and over), presented as a new patient with shoulder pain.

Intervention: There was no intervention related to the study. All the investigations, treatments and follow-ups were part of patient care for their shoulder problems.

Comparison: The researcher's estimations of diagnoses with the established colour-coded shoulder pain patterns

Outcome measures: To measure the accuracy, sensitivity-specificity of the previously established colour-coded shoulder pain maps and to establish an algorithm to establish an algorithm to facilitate the use of maps in clinical practice.

Phase Three

The aim of the third phase was to test intertester-reliability of the shoulder pain maps. This was achieved by three experience clinician as raters / testers, who were blind to the diagnoses, coding their two estimations of diagnoses for each pain map and using the algorithm and colour-coded shoulder pain patterns from the second phase of the study. Another aim in the third phase was to examine if the mapping system is easy to use or requires a lot of training. The testers' estimations were tested with Kappa values for their agreements.

Method:

It was conducted with a blinding method; the three testers were blind to the diagnoses. They were given algorithm, colour-coded maps for shoulder pain diagnoses. They were asked to code their two estimations on separate forms for each patient.

Design:

The third phase of the study was to analyse inter-tester reliability of the shoulder pain maps. Inter-rater or Inter-observer reliability is used to assess the degree to which different observers give consistent estimates of the same phenomenon (Trochim, 2006). Intertester reliability test was used in several studies and one of them was conducted with two hundred and fifty-five patients related to shoulder pain by Carter et al (2012). The current study's inclusion criteria of patients were similar to that one. They included the patients with shoulder pain arising within the gleno-humeral or associated joints. They excluded the patients with previous shoulder surgery, complex problems and concurrent cervical

pain and/or radiculopathy. They used percentage agreement and Cohen's kappa coefficient as in our study.

Participants:

Three senior clinicians participated to the third phase of the study as tester / rater. One of the clinicians was a senior physiotherapist who had over 20 years of experience in sport medicine with a PhD degree, the other clinician was an experienced post-FRCS (Ortho) orthopaedic doctor and working in an upper limp unit, and third one was a consultant radiologist with musculoskeletal interest.

Procedure:

A brief explanation was given to each participant about the study and the previously established algorithm and colour-coded pain maps for shoulder pain diagnoses from the second phase.

They were asked to code two estimations of diagnoses on a separate form for each patient of 194 from the second study using the algorithm and guidelines. The testers' first estimations were coded according to whichever group of shoulder diseases the estimations were belong to and accordingly given a number of 1 or 2 (as below- next page).

Patient No:

1.Estimation

☐ ACJ pathology ☐ Instability ☐ Calcific Tendinitis

☐ Rotator Cuff tear ☐ Frozen Shoulder ☐ Impingement ☐ GHJ Arthritis

2.Estimation

☐ ACJ pathology ☐ Instability ☐ Calcific Tendinitis

☐ Rotator Cuff tear ☐ Frozen Shoulder ☐ Impingement ☐ GHJ Arthritis

Data analysis:

The data from each of three participant was collected and converted to an excel sheet. The diagnoses codes were correlated with actual groups of diagnoses to find out accuracy of testers' estimation. Moreover, the group codes were used to calculate the agreement in between the testers (Group 1 and group 2). Later, the best of both estimations of testers were used to calculate the accuracy on specific disease and again to calculate the agreement between the testers for specific diseases. This was discussed with the study supervisor and the level of agreement was tested with Kappa statistical test. These were analysed with Kappa values to observe a score of how much consensus or homogeneity there was for the algorithm of the shoulder pain maps. Kappa is typically is used to look at how accurately a test can be repeated.

The excel sheet from the participants were transferred to statistics package (StatsDirect). Kappa co-efficient and inter-reliability tests were used to analyse this. Kappa (κ) of 0.5 or more is considered as good agreement and it is accepted of clinical significance. A value of 0.7 shows very good agreement. A κ of 1 means that there is perfect agreement (Harris & Taylor, 2008) (please refer to statistical terminology). The analysis was done using a

statistics package (StatsDirect). Opinion was sought from the local statistician for statistical analysis.

Ethical Implications:

This step did not include any additional treatment or investigations. The testers gave their verbal informed consent to contribute / join to the study. The participants had no financial interest in any product or service related to the study.

In summary- Phase three

Population: Adult patients (16 years and over), presented as a new patient with shoulder pain.

Intervention: There was no intervention

Comparison: To test the agreement between the testers using inter-tester reliability test

Outcome measures: Their estimations were tested with Kappa values for their agreements to observe a score of how much consensus or homogeneity there was for the algorithm of the shoulder pain maps.

The aim of the study was to fill the gap in assessing patients as a whole for common shoulder disorders, to develop a pathway from a subjective experience of patients to diagnosis of the pathology, This prospective blinded study with the methods of the three-phase gave a chance to study and analyse shoulder pain mapping extensively and vigorously.

Key Statistical Terminology

Analysis of variance (ANOVA): this is a group of statistical techniques used to compare the means of two or more sample to see whether they come from the same population (Harris & Taylor, 2008).

Blinding: A clinical trial design strategy in which one or more parties involved with the trial, such as the investigator or participant, who do not know which participants have been assigned to which interventions. Types of masking include none open label, single blind masking, double and blind masking (Clinical Trials, 2012). The opposite of a blinded study is described as an open label study (National Cancer Institute, 2012).

Confidence intervals (CI): typically used when instead of simply planning the mean value of a sample, if it is a range that is likely to contain the true population value (Harris & Taylor, 2008).

Kappa: a comparison of how well tests agree. Kappa value can vary from zero to 1 and 1 means that there is perfect agreement. 0.5 or more is considered a good agreement, a value of 0.7 shows very good agreement (Harris & Taylor, 2008).

Positive predictive value: the proportion of patients with positive test results who are correctly diagnosed (Altman and Bland, 1994a).

Negative predictive value: the proportion of patients with negative test results who are correctly diagnosed (Altman and Bland, 1994a).

Sensitivity: the proportion of true positives that are correctly identified by the test (Altman and Bland, 1994b).

Specificity: the proportion of true negatives that are correctly identified by the test (Altman and Bland, 1994b).

Results

The study was conducted in three phases; each phase had a different aim and was evaluated separately. The first phase aimed to ascertain different pain patterns for common shoulder disorders. It was a prospective study and Statistical Package for the Social Science (SPSS) 16 APA format was used to analyse demographic data and also to differentiate the pain distribution / radiation (the number of the affected cells/area) in the shoulder as well as the severity of the pain for each disease. However, the pain patterns were established manually from the excel sheets.

The second phase of the study aimed to test accuracy of the established colour-coded shoulder pain patterns and once the final diagnoses were obtained, they were correlated with the estimated diagnoses. VassarStats (Lowry, 2013) was used for sensitivity and specificity and StatsDirect (2013) package was used for analysis of the other data such as age, agreement between estimation and diagnoses, agreement between map group and disease group. At the end of the second study, an algorithm on how-to-read shoulder pain maps was created.

The third phase of the study aimed to analyse inter-tester reliability of the shoulder pain maps for three clinicians' agreement level and the data was analysed with StatsDirect. The clinicians used the algorithm and the established colour-coded shoulder pain maps to test inter-tester reliability.

All the reports of statistical analysis were added to the appendices.

Phase One

The first phase of the study included six shoulder disorders according to the adequacy of the number of patients for each disease once the results were analysed. The included disorders were acromioclavicular joint pathology, instability: Bankart's, SLAP etc, calcific tendonitis, rotator cuff pathology, impingement syndrome, gleno-humeral joint arthritis.

There was a distinct age difference between some of the disorders. For example, the patients with instability were generally younger than the other groups. The mean age of the instability group was 34.4 years (see table ii). This was followed by calcific tendonitis with 46.5 years. The oldest age group was gleno-humeral arthritis with an average of 69.8 (table ii).

Mean age- Phase one:

Disorder	1-ACJ	2-Instability	3-Calc ten.	4-rot c.	6-imp	7-GHJ
Mean Age 1.phase	58.64	34.38	46.5	66.04	57.64	69.83

Table ii: Mean ages for each disorder (Phase 1)

(1-ACJ: Acromioclavicular joint pathology, 2-Instability: Bankart's, SLAP etc, 3-Calc ten.: Calcific tendonitis, 4-rot c.: Rotator cuff pathology, 6- imp: Impingement syndrome, 7- GHJ: Gleno-humeral joint arthritis)

Note: number 5 was not included to the table to avoid any confusion as number 5 is frozen shoulder when the results are combined with the second phase of the study.

Shoulder pain patterns and distributions for each shoulder disorder were obtained by observing and checking the excel data sheet manually in the first phase of the study and the summary is shown on the table below (see table iii). (Please note that number 5 is frozen shoulder and it is empty on the table as it was not included in the first phase.) The table content was converted to colour-coded shoulder maps at the end of the first phase (see figures II-i, ii, iii, iv-a, vi and vii in appendices).

The shoulder pain distribution table according to the pain types, region and disorders from the first phase of the study:

Types of pain and regions: Overall: both anterior&posterior	Predominant Pain types around shoulder N:				Predominant Pain types around arm N:				Predominant Pain types below elbow N:			
Diagnosis	S	D	B	P	S	D	B	P	S	D	B	P
1	12	1	1			5				2		
2	8	9	1		4	4		1				
3	5	1			4		1					
4	15	7			3	11		1	3	8		1
5 **												
6	18	6	4		8	16	2	1	3	9		7*
7	3	2	1		3	2	1		3	2		

Table iii: Types and radiation of shoulder pain- phase one (Bayam et al., 2011)

(N: number of the patients with type of pain description, 1-Acromioclavicular joint pathology, 2-Instability: Bankart's, SLAP etc, 3- Calcific tendonitis, 4-Rotator cuff pathology, 5-Frozen shoulder, 6- Impingement syndrome, 7-Gleno-humeral joint arthritis,S: sharp,shooting or stabbing pain, D: dull or aching pain, B: burning pain, P: pins&needles / numbness)

* 7 patients with impingement syndrome had described pins & needles sensation in their hands, mainly on the dorsum. These patients were assessed to exclude particularly other distal upper limb problems for example carpal tunnel syndrome and other nerve compression disorders.

**Frozen shoulder (5) was not included in the first study as one of the diagnosis.

Character and Radiation of Pain for Common Shoulder Disorders

Pain from ACJ pathology was predominantly localised to the anterior aspect of the shoulder and was sharp and stabbing in nature. Pain did not radiate down the forearm. The area of pain distribution was smallest among all the disorders (figure II-i, table iii).

The patients with shoulder instability including SLAP (Superior Labrum from Anterior to Posterior) and Bankart`s lesions had predominately localised pain around the shoulder. It was a mixture of sharp and dull pain in character without radiation to the forearm or even not to the elbow (figure II-ii, table iii). They were younger than the other disease groups (table ii).

The patients with calcific tendonitis described predominantly a sharp / shooting pain in the region of the shoulder with no radiation to the elbow or hand (figure II-iii, table iii).

Patients with rotator cuff tears demonstrated sharp pain around the shoulder and radiating dull / aching pain down to the elbow and the forearm (figure II-iv-a, table iii).

The patients with impingement syndrome showed sharp pain around the shoulder and radiating dull / aching pain down to the forearm. It was similar to rotator cuff pattern, however, there was pins & needles / numbness around the hand (figure II-vi, table iii).

In GHJ arthritis, a mixed pattern of pain was described that was sharp, shooting, burning and dull aching in nature and affected nearly the entire arm, from shoulder to hand (figure II-vii, table iii).

Burning sensation was rarely described. When described, this was mainly by patients who had impingement and GHJ arthritis (table iii).

The Distribution Area of Pain and the Severity (First Phase Study):

Using SPSS 16.0, statistical package, P-P plot for distribution showed that both the severity of the pain and the number of the areas were normally distributed. On analysis of the VAS scores, GHJ arthritis patients had the most severe pain with a mean score of 7.83 / 10, closely followed by patients with impingement with a mean score of 7.80. The least severe pain was described by the patients with ACJ pathology, the mean was 6.43 / 10. Similarly, the value for instability patients was 6.72 / 10. The remaining VAS pain severity scores were 7.05 and 7.50 for rotator cuff tears and calcific tendinitis respectively (table iv). Statistically, there was not a significant difference in the range of pain between the diagnoses (F (5, 87)= 1.138, p=0.347).

The pain map consisted of a total of 28 marked sections or cells, 14 on each side of the arm (figure I). Analysis demonstrated how well the pain localised in a specific area for each particular shoulder condition (table iv). This ranged between 3.86 for ACJ pathology to 12.5 for gleno-humeral arthritis, indicating the diffuse nature of pain from GHJ arthritis and the pinpoint nature of pain from ACJ pathology.

The table for distribution area of pain and the severity:

Diagnosis	1	2	3	4	6	7
Number of cells (mean)	3.86	6.88	5.5	5.60	7.21	12.5
Severity of pain (VAS)	6.43	6.72	7.5	7.05	7.80	7.83

Table iv: 1- The mean of total number of cells marked by the patients (regardless the type of pain) (in total, there were 28 cells anteriorly&posteriorly on the map) and 2- the severity of the pain (VAS) for each disease.

(1-Acromioclavicular joint pathology, 2-Instability: Bankart's, SLAP etc, 3- Calcific tendonitis, 4-Rotator cuff pathology, 6- Impingement syndrome, 7-Gleno-humeral joint arthritis)
Note: There is no number 5 on the table. In the combined results, number 5 is frozen shoulder. To avoid any confusion, number 5 was kept empty on the table

Whilst Pearson's correlation tests demonstrated statistically weak correlation between the number of the areas and severity of the pain ($r=0.194$, $p= 0.64$), however, one-way ANOVA showed a meaningful difference in the number of areas marked by the patients between the groups (F (5, 87) = 3.550, $p=0.006$). Post hoc multiple comparisons showed meaningful differences were especially between GHJ and rotator cuff tears ($p= 0.014$) and also, between GHJ arthritis and ACJ pathology ($p= 0.002$).

At the end of Phase one, all the colour-coded maps were established for the above six shoulder disorders (figure II i, ii, iii, iv-a, vi and vii) and these were used to be tested in the second phase of the study.

Phase Two

In this phase, the final diagnoses were correlated with the estimated diagnoses. One aim of the second phase was to recruit many more patients to the study in order to test the colour-coded shoulder pain maps and improve these if required. Despite the initial inclusion and exclusion criteria, many more patients (128) were included initially but, excluded from the study at later stage. 34 of them were excluded due to inappropriate marking such as using different marks, illegible marks, or not marking the forms at all, although they all have been explained about how to fill the forms. 30 of them were excluded due to insufficient clinical information or documentation about their diagnoses (probably related to drop out or loss to follow-up). 14 of them were excluded due to previous shoulder operation on the same shoulder. 25 of them were excluded due to multiple shoulder problem/diagnoses. And 25 of them were excluded due to irrelevant/different diagnoses, which were not within the framework of this study such as metastatic disease or neurological disorders as well as neck problems which were missed during the initial exclusion. 194 patients were included in the second study.

Mean Age- Phase Two

The mean age of the instability group was 28.8 years in the second study (see table v). This was followed by calcific tendonitis with 43 years. The oldest age group was gleno-humeral arthritis with an average of 65.5 (table v). These results were similar to phase one results (see above, table ii).

Mean age- Phase two:

Disorder	1- ACJ	2- Instability	3-Calc ten.	4-rot c.	5- frozen	6-imp	7- GHJ
Mean Age 2. phase	48.4	28.8	43.0	53.4	51.8	49.5	65.5

Table v: Mean ages for each disorder (Phase two)

(1-ACJ: Acromioclavicular joint pathology, 2-Instability: Bankart's, SLAP etc, 3-Calc ten.: Calcific tendonitis, 4-rot c.: Rotator cuff pathology, 5-frozen: Frozen shoulder, 6-imp: Impingement syndrome, 7- GHJ: Gleno-humeral joint arthritis)

All the patients (194) had clinical examination, antero-posterior, axillary and supraspinatus outlet views of the affected shoulder and shoulder ultrasound by a senior upper limb surgeon as part of their clinical assessments. In addition to the initial investigations, which were x-rays and ultrasound scans, on the day of first assessment, 55 patients had MR arthrogram, 39 patients had MR scan and 10 patients had CT arthrogram later. In total, 117 patients had further scans / investigations (table vi).

The table for further scans- phase two:

Further Scans (117 out of 194 in total)			
MR scan	MR Arthrogram	CT Arthrogram	Shoulder Arthrography
39	55	10	13

Table vi: Further imaging in addition to the investigations on the first day

130 patients out of 194 had surgery, 14 patients had arthrographic hydrodilatation, 9 patients had injection and physiotherapy, 14 patients had physiotherapy, 21 injection and 3 ultrasound guided removal of calcium deposits (barbotage). In total, 191 out of 194 patients in the second study had further treatment for their shoulder disorders (table vii).

The table of further managements for patients' shoulder disorders

Operation	Hydrodilation	PhysiotherpyAnd injection	Physiotherapy	injection	u/s guidance removal of ca deposits	No further treatment
130	14	9	14	21	3	3

Table vii: Treatments of patients with shoulder disorders- phase two

In the second phase, the diagnoses were established after completion of all the investigations, assessments, treatments, operations and follow-ups.

Testing Colour-coded Shoulder Pain Maps

The accuracy of the first diagnoses for individual disorders was 45.4% (88 correct diagnoses out of 194) and when including the second estimations, the accuracy was 62.4 % (121 out of 194). 22 out of 194 patients were diagnosed with frozen shoulder, which were included in those percentages.

Prediction of researcher on diagnoses of shoulder disorders using colour-coded shoulder pain maps:

Diagnosis	1st estimated diagnosis	2nd estimated diagnosis	TOTAL Correct diagnosis	Not diagnosed	Total patients
1-ACJ path.	13	12	25 (86.2%)	4	29
2-Instability	47	4	51 (89.5%)	6	57
3-Calcific tend.	2	3	5 (71.4%)	2	7
4-Rotator cuff	5	3	8 (38.1%)	21	29
5-Frozen shoulder**	0**	0**	0**	22**	22
6-Impingement	20	10	30 (65.2%)	16	46
7-GHJ arthritis	1	1	2 (50%)	2	4
Total	88	33	121 (62.4%)	73	194

Table viii: Prediction of diagnoses by the researcher

(**Note: 22 of the patients were diagnosed with frozen shoulder, which did not have an established pattern/map in the first phase of the study)

There was a high prediction rate for instability and ACJ pathology and it was low for rotator cuff pathology (table viii).

The table for sensitivity and specificity of estimations using shoulder pain maps:

Diagnosis	Sensitivity	Specificity
1-ACJ path.	0.62	0.97
2-Instability	0.78	0.85
3-Calcific tend.	0.5	0.99
4-Rotator cuff	0.33	0.95
6-Impingement	0.67	0.93
7-GHJ arthritis	0.33	0.99

Table ix: Sensitivity and specificity of estimations for each shoulder disorders

According to the above table (ix), in the study, the estimations for the diagnoses of the maps were most sensitive for instability. The patients with instability had mostly shown and marked the same pattern for their pain. This was followed by impingement and ACJ pathology. It was least sensitive for rotator cuff and GHJ arthritis. However, there were a very small number of the patients with GHJ arthritis.

Table x, xi and xii were used to evaluate the pain patterns of common shoulder disorders and these were combined with the results from the first phase to describe and improve the colour-coded patterns. Besides, these tables helped establishing pain mapping algorithm.

The Pain distribution tables from the second phase of the study:

Types of pain and regions: Overall results	Pain around shoulder N:	Pain around upper arm N:	Pain around elbow N:	Pain below elbow N:
Diagnosis →N				
1→29	28 95.6%	8 27.6%	2 6.9%	0
2→57	100%	26 45.6%	3 5.3%	1 1.8%
3→7	100%	4 57.1%	1 14.3%	0
4→ 29	100%	21 72.4%	8 27.6%	4 13.8%
5→22	21 95.5%	20 90.9%	13 59%	8 36.4%
6→46	100%	40 87%	29 63%	20 43.5%
7→4	100%	100%	3 75%	3 75%

Table x: Overall shoulder pain distribution (without pain type description) – second phase of the study

(N: number of the patients, 1-Acromioclavicular joint pathology, 2-Instability: Bankart's, SLAP etc, 3- Calcific tendonitis, 4-Rotator cuff pathology, 5-Frozen shoulder, 6-Impingement syndrome, 7-Gleno-humeral joint arthritis, S: sharp, shooting or stabbing pain, D: dull or aching pain, B: burning pain, P: pins & needles / numbness)

The summary table of pain types and distribution anteriorly for each shoulder disorder- phase two:

Types of pain and regions: Overall: anterior	Predominant Pain types around shoulder N:				Predominant Pain types around arm N:				Predominant Pain types below elbow N:			
Diagnosis → N	S	D	B		S	D	B		S	D		P
1→29	20 69%	5 17.4 % 6*	2*		2 6.9%	6 20.7 %				2 6.9%		1 3.4%
2→57	43 75.4%	12 21% 34*	6*		6 10.5 %	13 22.8 % 3*	2 1*					5 8.8%
3→7	5 71.4%	1 14.3 % 4*	1*		1 14.3 %	2 28.6 %				1 14.3 %		1 14.3%
4→ 29	20 69%	8 27.6 % 7*	1 3.45 % 10*		10 34.5 %	7 24.1 % 3*	2 6.9 %		1 3.4 %	5 17.2 %		5 17.2%
5→22	17 77.3%	2 9% 5*	4*		6 27.3 %	11 50 %	1 4.5 %		3 13.6 %	7 31.8 %		3 13.6%
6→46	32 70%	8 17.4 %	1 2.17 %		10 21.7 %	26 56.5 %	1		1	20 43.5 %		13 28.3%
7→4	4 100%	2*	1*		3 75%	1*	1*		2 50 %	1 25% 1*		1 25%

*= non-dominant pain in that region

Table xi: Types of shoulder pain and regions anteriorly, second phase

(N: number of the patients, *1*-Acromioclavicular joint pathology, 2-Instability: Bankart's, SLAP etc, 3- Calcific tendonitis, 4-Rotator cuff pathology, 5-Frozen shoulder, 6-Impingement syndrome, 7-Gleno-humeral joint arthritis, S: sharp, shooting or stabbing pain, D: dull or aching pain, B: burning pain, P: pins & needles / numbness)

The summary table of pain types and distribution posteriorly for each shoulder disorder- phase two:

Types of pain and regions: Overall: posterior Diagnosis→ N	Predominant Pain types around shoulder N:				Predominant Pain types around arm N:				Predominant Pain types below elbow N:			
	S	D	B	P	S	D	B	P	S	D	B	P
1→29	12 41.4 %	5 17. 25 %				2 6.9 %						
2→57	25 43.9 %	13 22. 8 % 23 *	1 1. 8 % 6*			6 10. 5 % 2*				3 5. 3 %		4 7%
3→7	5 71.4 %	1*				1 14. 3%						
4→ 29	9 31%	11 38 %	1 3. 4 % 5*			4 13. 8%	1 *			3 10 %		1 3.4 %
5→22	14 63.6 %	2 9 %	1 4. 5 %			7 31. 8%			1 4. 5 %	8 36 .4 %		2 9%
6→46	20 43.5 %	12 26 % 16 *	6*			17 40 % 3*	1 2.2 % 1*			16 34 .8 % 1*		8 17. 4%
7→4	3 75%	1*				1*			1 25 %	1 25 %		1 25 %

The table continues to the next page

*= non-dominant pain in that region

Table xii: Types of shoulder pain and regions posteriorly, second phase

(N: number of the patients, *1*-Acromioclavicular joint pathology, 2-Instability: Bankart's, SLAP etc, 3- Calcific tendonitis, 4-Rotator cuff pathology, 5-Frozen shoulder, 6-Impingement syndrome, 7-Gleno-humeral joint arthritis, S: sharp, shooting or stabbing pain, D: dull or aching pain, B: burning pain, P: pins & needles / numbness)

Pins & needles or numbness in the hands and wrists were mostly described by patients with impingement syndrome in both studies [25% (7/28) and 30 % (14/46) in the first and second phases study, respectively]. Despite high percentage of pins & needles in GHJ arthritis, the total number of patients with the disorders was very low (table xiii).

Numbness in the hand and wrist in the second phase of the study:

Diagnosis (N)	Number of patient with numbness / pins & needles in the hand (N)	Total Percentage of numbness or pins & needles in the hand and wrist
1- ACJ pathology (29)	2 in total (out of 29) 2 front 0 back	6.9%
2- Instability (57)	7 in total (out of 57) 5 front 4 back	12.3%
3- Calcific Tendinitis(7)	1 in total (out of 7) 1 front 0 back	14.3%
4- Rotator cuff (29)	4 in total (out of 29) 4 front 1 back	13.8%
5- Frozen shoulder(22)	3 in total (out of 22) 3 front 2 back	13.6%
6- Impingement (46)	14 (out of 46) 11 front 8 back	30.4%
7- GHJ arthritis (4)	2 in total (out of 4) 1 front 1 back	50% (small number of patients)

Table xiii: Numbness / pins & needles in the hand and wrist

(N: number of the patients)

In the light of the results in testing the colour-coded shoulder maps, the second step phase had revealed mainly a problem with rotator cuff pathology in the initial pain patterns. Therefore, it was analysed again using the combination of maps from both the first and second phases, which led to some modification in its description and pattern which were slightly changed accordingly. Following the analysis of the second study phase, another disorder, frozen shoulder was also added as a new pattern. The colour-coded mapping of frozen shoulder was produced in addition to the previous six shoulder disorders which were included in the first phase.

Additional Changes to Phase One Results in Distribution and Character of Pain in Shoulder Disorders

As the testing of phase one results revealed some problems with rotator cuff pattern, the description about rotator cuff changed using the combined results of shoulder pain maps from both phase one and two studies. Also, the paragraph on impingement syndrome changed as below:

Patients with rotator cuff tears demonstrated sharp pain around the shoulder and radiating dull / aching pain towards the elbow. Mostly, the pain was limited to around the elbow and was not radiating down to the forearm or hand (figure II-iv-b, table x, xi, xii).

The patients with impingement syndrome showed sharp pain around the shoulder and radiating dull / aching pain down to the forearm. It was similar to rotator cuff pattern,

however, there was further radiation of the pain below elbow as well as pins & needles / numbness around the hand (figure II-vi, table x, xi, xii).

Frozen shoulder showed a pattern of mostly sharp widespread pain around the shoulder mixed with burning pain and radiation to elbow which was distributed posteriorly around the arm down to elbow level but generally no radiation below the elbow level (figure II-v, table x, xi, xii).

Burning sensation was described by the patients who were diagnosed with frozen shoulder in addition to the patients with impingement and GHJ arthritis even though this pain type was rarely marked on the map by the patients (table iii, x, xi, xii).

At the end of the second phase, all tables from phase one and two were combined and common shoulder disorders were classified according to the pain patterns, character and distribution as explained in the next section. The age was added to the algorithm for group one (the pain localised around the shoulder, see algorithm) but for group two, the age did not contribute to the differential. The severity of the pain was not included because this did not show a significant statistical difference between the groups in the first phase of study.

Classification of Common Shoulder Disorders According to Pain Distribution / Radiation

Since both phase one and some part of phase two looked into the same details on patients' experience to diagnoses of shoulder pathologies and the results of two phases were combined here to classify the pain patterns for common shoulder disorders. Seven common shoulder disorders were identified on the patients who presented to the shoulder clinic and filled the custom-made shoulder pain maps. According to the pain distributions of the seven shoulder disorders, they were divided into two groups based on the distribution of the pain area whether localised to shoulder or not. They were called as group 1 and group 2; group 1 represented the disorders with shoulder pain localised only around shoulder (ACJ pathology, Calcific tendinitis, and instability) and group 2 represented the disorders that the pain radiates beyond shoulder (rotator cuff pathology, frozen shoulder, impingement, and GHJ arthritis). Later, the second group was also divided into two subgroups based on further radiation of the pain.

Classification of Common Shoulder Disorders

Group 1: Pain localised around shoulder

Disease 1: Acromioclavicular (ACJ) pathology

2: Instability (Bankart's, SLAP etc)

3: Calcific tendonitis

Group 2: Pain radiating beyond shoulder

A- To around elbow

4: Rotator cuff pathology

5: Frozen shoulder

B- Beyond elbow-down to the hand

6: Impingement syndrome

7: Gleno-humeral joint (GHJ) arthritis

When the diseases were divided into two groups (group 1 and 2) according to the pain distribution (group 1: localized to the shoulder, group 2: radiation beyond shoulder) as per the classification above, the agreement between map group and disease group was analysed. It showed a substantial agreement between the two groups (81.96%) (Observed agreement = 81.96%, Expected agreement = 50.28%, Kappa (κ) = 0.64, 95% confidence interval = 0.50 to 0.78, $P < 0.0001$).

The sensitivity of maps for group match was 0.76 for group 1 and 0.87 for group 2. The specificities were 0.87 and 0.76 for group 1 and group 2, respectively (table xiv).

The table for estimation of disease groups matching shoulder pain maps:

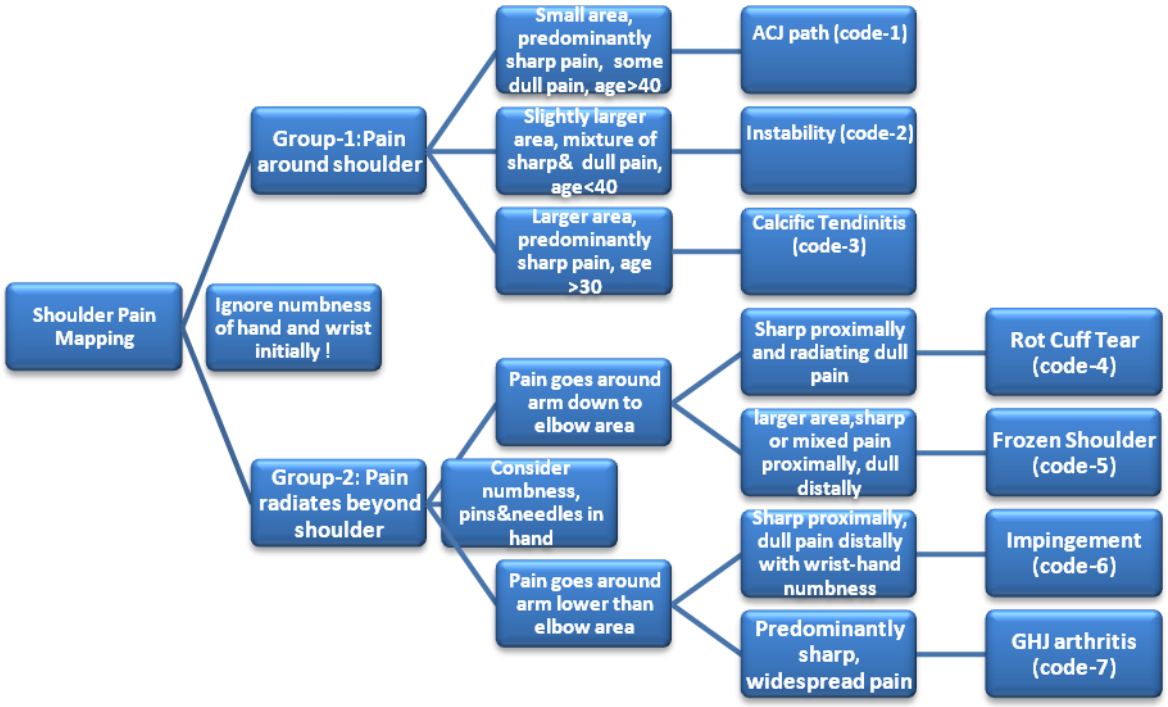
Estimation of Disease Group	Sensitivity	Specificity
Group 1 diseases (ACJ path, Instability, Calc. tend.)	0.76	0.87
Group 2 diseases (Rot.Cuff. frozen, impingement, GHJ arth.)	0.87	0.76

Table xiv: Estimation of Disease Group, sensitivity and specificity

(ACJ path: Acromioclavicular joint pathology, 2-Instability: Bankart's, SLAP etc, 3-Calc tend.: Calcific tendonitis, 4-rot cuff.: Rotator cuff pathology, 5-frozen: Frozen shoulder, 6-impingement: Impingement syndrome, 7- GHJ arth: Gleno-humeral joint arthritis)

Following the analysis of the second step of the study, an algorithm is formed as below. This aimed to provide guidance about how to read the shoulder pain maps as well as to assist facilitating the third phase of the study for inter-tester reliability.

Shoulder Pain Mapping Algorithm



The extension of the marked area of pain from less to more (the numbers shows the disorder codes):

1 < 2 < 3 < 4 < 5 < 6 < 7

Figure III: Shoulder Pain Mapping Algorithm

Phase three:

In the third phase, all three testers / raters gave two estimations for each map of 194 patients using the algorithm and the established colour-coded pain maps for shoulder disorders and their best estimations were used for statistical analysis; estimation of diagnoses, agreement between testers / raters about these estimations, sensitivity and specificity for individual shoulder disorders. However, only their first estimations were used to analyse the disease group; sensitivity and specificity for disease groups and agreement between testers / raters.

Estimating the Disease Group:

There were high estimation ratios for the disease groups (group 1 and 2) of shoulder in the raters' first estimation. Rater A had an estimation of 75.26%, whilst Rater B and Rater C had 81.44% and 80.93 respectively (table xv).

The table of estimations by testers:

Raters / testers	Expected agreement	Observed agreement	P value
A	49.87%	75.26%	P < 0.0001
B	50%	81.44%	P < 0.0001
C	50.23%	80.93%	P < 0.0001

Table xv: Estimation of Disease Groups by Three Testers

Furthermore, the results showed high sensitivity and specificity for the disease group, which was obtained from the estimations of the testers.

The raters' estimations about the disease groups:

Disease Group	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
	Rater A	Rater A	Rater B	Rater B	Rater C	Rater C
Group-1	0.80	0.71	0.83	0.80	0.76	0.85
Group-2	0.71	0.80	0.80	0.83	0.85	0.76

Table xvi: Sensitivity and specificity for the disease groups

The table xvii shows a substantial agreement between the raters on the estimation (first estimation) of disease groups.

The table of agreement between raters:

	Expected agreement	Observed agreement	Kappa (κ)	P
Between 3 raters			0.70	P < 0.0001
Between rater A and B	50%	82.5%	0.65	P < 0.0001
Between rater A and C	49.65%	81.96%	0.64	P < 0.0001
Between rater B and C	50%	90.21%	0.80	P < 0.0001

Table xvii: Agreement of the raters on the disease groups

Estimating the Diseases Individually:

The matching of best estimations of three raters with the final diagnoses of individual shoulder disorders were from 58 to 64 %.

The table of estimations of raters on the individual shoulder disorders:

Raters / testers	Expected agreement	Observed agreement	P value
A	19.41%	60.82%	P < 0.0001
B	19.39%	63.92%	P < 0.0001
C	17.35%	58.25%	P < 0.0001

Table xviii: Best estimation of individual shoulder disease by raters

Sensitivity shows the proportion of the patients with the disease, whose tests are positive and specificity shows the probability that the patients do not have the disease when they are disease free (Faragher, 2005). The table shows constantly higher sensitivity of map estimation for instability and also better results for rotator cuff disorders after the improvement following the second step of the study (table xix).

The table of sensitivity and specificity of testers for individual shoulder disorders:

Diagnosis	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
	Rater A	Rater A	Rater B	Rater B	Rater C	Rater C
1-ACJ path.	0.55	0.90	0.51	0.96	0.55	0.93
2-Instability	0.74	0.89	0.81	0.97	0.65	0.96
3-Calcific tend.	1	0.97	0.43	0.99	0.71	0.95
4-Rotator cuff	0.55	0.87	0.72	0.88	0.52	0.84
5-Frozen Shoulder	0.55	0.95	0.32	0.94	0.55	0.95
6-Impinge-ment	0.5	0.93	0.54	0.92	0.52	0.95
7-GHJ arthritis	0.5	0.99	0.25	0.97	1	0.94

Table xix: Sensitivity and Specificity Results of Testers for Individual Shoulder Disorders

(ACJ path: Acromioclavicular joint pathology, 2-Instability: Bankart's, SLAP etc, 3-Calcific tend.: Calcific tendonitis, 4-rotator cuff.: Rotator cuff pathology, 5-frozen: Frozen shoulder, 6- impingement: Impingement syndrome, 7- GHJ arthritis: Gleno-humeral joint arthritis)

There were also constantly similar sensitivity values for ACJ pathology and impingement syndrome.

The agreement of estimation between 3 testers who gave 2 estimations for each map was calculated and their best estimation was used for statistical analysis. It showed a substantial agreement between the raters (Kappa (κ) = 0.71) (table xx).

The table of overall agreements between the raters:

Disease	Kappa(κ)
1-ACJ path.	0.77
2-Instability	0.73
3-Calcific tend.	0.62
4-Rotator cuff	0.66
5-Frozen Shoulder	0.67
6-Impingement	0.77
7-GHJ arthritis	0.53
Overall	0.71

Table xx: General agreement over all (7) categories with 3 raters per disease

The agreement of estimations for individual diseases between raters (A and B, B and C, A and C) was very good (table xxi).

The table of agreement between raters on individual shoulder disorders:

	Expected agreement	Observed agreement	Kappa	P
Between rater A and B	19.13%	88.14%	0.85	P < 0.0001
Between rater A and C	17.59%	82.47%	0.79	P < 0.0001
Between rater B and C	19.17%	93.78%	0.92	P < 0.0001

Table xxi: Raters agreement on the estimations of the individual shoulder disorders

In summary, the results section showed establishing colour-coded shoulder pain maps in the first phase, testing these, improving them and establishing an algorithm about how-to-read shoulder pain maps in the second phase and assessing the inter-tester reliability by raters for both the disease groups and individual shoulder disorders in the third phase. These results were discussed in the next section in the light of existing literature.

Discussion:

This chapter discussed and analysed the results of the three phases in light of the existing published studies. It included a summary of the study and results, discussion on pain distribution, sensitivity - specificity - reliability, pain maps, limitation and conclusion.

Pain mapping for specific disorders was described in the literature for face, back and hip pain, but not for shoulder pain. Literature search revealed a gap in this area. The aim of the study was to fill the gap in assessing patients as a whole for common shoulder disorders, to develop a pathway from a subjective experience of patients to diagnosis of the pathology, to ascertain specific patterns of pain in patients with common shoulder disorders, to describe comprehensive shoulder pain maps and to test these.

The study was conducted in three phases to achieve the above aims. A prospective blinded study was used in the first phase to achieve establishing pain patterns for common shoulder disorders and six specific patterns of shoulder diseases were described. The colour-coded pain maps were acromioclavicular joint pathology, instability: Bankart's, SLAP etc, calcific tendonitis, rotator cuff pathology, impingement syndrome, gleno-humeral joint arthritis. The first phase included 94 patients.

Although the pain severity was not statistically significant, there was a difference especially between ACJ pathology, instability and GHJ arthritis. Trend for severity of the pain showed lower VAS for ACJ and higher VAS for GHJ arthritis. Although it was not statistically significant, with greater number, the trend may be important however, equally it

may remain not important. On analysis of the pain distribution, it indicated the diffuse nature of pain from GHJ arthritis and the pinpoint nature of pain from ACJ pathology.

The second phase aimed to test these maps with another prospective blinded study and with a larger group of patients. The accuracy of mappings was tested. The provisional / estimated diagnoses were marked immediately by the researcher, who was blind to the diagnoses, after each shoulder map marked by the patients. However, the coding of the diagnoses for those patients in the second study was delayed until all the operations and postoperative outpatient follow-ups were completed in order to obtain the final and definitive diagnoses. The second study included 194 patients.

The second phase showed high sensitivity especially for instability as well as good sensitivity for ACJ pathology and impingement syndrome. The results for rotator cuff pathology were poor initially; however, the description of its pattern had been improved accordingly. Another shoulder disorder was added to previous six diagnoses, which was frozen shoulder.

At the end of the second phase, the previously established shoulder pain patterns were improved. Two main groups of diseases were described according to the pain distribution / radiation and the shoulder disorders were classified accordingly. There was good agreement between the map group (1 and 2) and the disease group (1 and 2). An algorithm was developed to guide the testers/ raters of the third phase on how-to-read the shoulder pain maps.

The third phase aimed to test inter-tester reliability with three testers and it achieved this with blinding of the testers to the diagnoses on a large group of patients (194 patients from

the second phase of the study). Three testers gave their estimations according to the algorithm and the established colour-coded shoulder pain maps. The results showed a substantial agreement between testers for both disease groups and the individual shoulder diseases, which indicates the reliability of the algorithm and established colour-coded maps. Likewise, sensitivity and specificity for both disease groups and individual shoulder disorders were tested and calculated.

The above summary of the study was discussed further as below.

Pain Distribution:

The current study showed specific pain patterns for common shoulder problems. Many different shoulder disorders cause similar symptoms and patterns of pain (Larson et al., 1996). In the literature, there were some studies which attempted to specify the pain patterns as well as the current study. Gerber, Galantay and Hersche (1998) aimed to ascertain the distribution of ACJ and subacromial impingement pain by injecting hypertonic saline into the ACJ and subacromial spaces of normal subjects. ACJ irritation resulted in "burning" pain felt over the joint, deep in the supraspinatus fossa and in the upper trapezius. This is similar to our findings in the mapping of ACJ pain, where pain was mainly localised to the anterior and dorsal aspect of the shoulder but differed in the distribution of pain down the posterior aspect of the arm. The pain was mainly stabbing in nature but also had dull and burning components.

The results from the current pain mapping study showed localised shoulder pain for certain shoulder conditions such as instability, calcific tendonitis and ACJ pathology. In instability, as well as sharp element of the pain, there was strong description of dull pain by the patients and majority of patients were younger than 40 years old in the current study. The pain distribution for calcific tendinitis in some literature was described as from the point of the shoulder to the deltoid insertion commonly and, less frequently, to the neck (Woodward, 2013), which is not much different than the current study's findings. In the current study, this was generally described as sharp pain by the patients, and occasionally with mixed nature of sharp and dull pain only around the shoulder area with no radiation below the shoulder. The nature of this pattern with wider area involvement

and mixture character of the pain was different from the ACJ pathology despite localized nature of the pain.

On the contrary, there were differences between the current study and previous literature in relation to radiation of the pain for other shoulder disorders. A far greater radiation of pain occurred in impingement syndrome and gleno-humeral arthritis in our study. Subacromial irritation in Gerber's paper resulted in an "intense" pain mainly in the lateral border of the acromion and the lateral portion of the deltoid muscle (Gerber et al., 1998). Dutton describes the pain due to rotator cuff pathology and impingement is usually felt over the anterior or lateral part of the shoulder and he mentions that this pain is characterised by radiation down the upper arm, and is aggravated with overhead activities. However, the pain that radiates beyond the elbow is far less likely to be due to shoulder pathology, particularly if it is associated with any sensory disturbance in the limb such as distal radiation or pain, numbness or paresthesia (Dutton, 2008). Similarly, Woodward and Best (2000) concluded that the pain related to impingement usually occurs over the anterolateral aspect of the shoulder, often with some radiation to, but not usually beyond, the elbow.

Besides, they mention that in frozen shoulder, the discomfort is localised near the deltoid insertion. Whereas, our two studies uniquely showed that at least two of the common shoulder diagnoses, impingement syndrome and gleno-humeral arthritis, show pain radiation beyond the elbow and some associated sensory disturbance. Furthermore, the current study showed that around 60% of the patients with frozen shoulder described their pain radiating down to elbow level. Frozen shoulder pattern was described in the second

phase of the study and there were some similarities to rotator cuff pathology pattern, yet, frozen shoulder pain was marked by the patients more widespread on the pain map than the pain distribution described for rotator cuff.

Pins & needles / numbness, which 30 % of the patients described around the hand in impingement syndrome also shows the extension of radiation of the pain in those patients. It is an interesting finding that the patients with impingement syndrome constantly described this symptom at a higher percentage than the other disorders. The high percentage of sensory disturbance may worth to look into more details and study as the other shoulder disorders did not show this pattern very often. We did not find further explanation or reason behind this pattern in the literature.

The precise location of pain is important in orthopaedics but does not always correlate with the site of pathology (Rodowsky & Bigliani, 1997). Despite the explanation that pain arising in deeper structures is more diffuse and sometimes has an unexpected distribution (Apley & Solomon, 1993), this needs to be highlighted and may require further explanation.

The other shoulder problem, GHJ arthritis pain had the most widespread distribution among all diagnoses despite the limitation of the study due to the very limited number of the patients with GHJ arthritis. The issue related to widespread distribution of GHJ arthritis was discuss further as below in pain maps section.

Sensitivity, Specificity, Reliability and Algorithm:

The second study showed high sensitivity and specificity for the disease groups. Sensitivity implies the proportion of the patients with the disease, whose tests are positive and specificity implies the probability that the patients does not have the disease when they are disease free (Faragher, 2005). There is no general agreement on what the acceptable levels of sensitivity and specificity for an assessment test are. Acceptable levels vary depending upon the intent of the test (Methodology tables, 2002). There were some previous studies with assessment of pain mapping for example back pain mapping by Pang et al. (1998) but they did not assess the sensitivities and specificities. So far, there were no other sensitivity and specificity studies on patients' subjective experience of shoulder pain or pain map as diagnostic tool in the literature, Therefore, there was no other study for comparison.

However, the literature described sensitivities and specificities for different clinical tests in the assessment of shoulder disorders as objective findings. Comparing to the meta-analysis of objective physical tests' results, we believe that our sensitivity and specificity results from subjective patients' experience to the diagnoses of pathology were generally acceptable. A meta-analysis showed the sensitivity and specificity of Neer test for impingement syndrome was 79% and 53%, respectively, and for the Hawkins-Kennedy test was 79% and 59%, respectively (Hegedus et al., 2008), whilst another study showed the sensitivity decreased to 40.3% when four tests of impingement syndrome were simultaneously positive, the specificity for the diagnosis was 98.5% (Fodor, Poanta, Felea,

Rednic, & Bolosiu 2009). In the current study, with self-completed patient mapping questionnaire, the sensitivity was high, and nearly as valuable as the other clinical tests (sensitivity 67.4%, the specificity 93.2%).

The clinical tests for superior labral antero-posterior (SLAP) tears showed lower sensitivity and specificity of the Speed test; 32% and 61%, respectively (Hegedus et al., 2008). In the current study, all the instability diseases (SLAP, Bankart's, etc.) are grouped as one and the sensitivity and specificity of the clinical tests were relatively high. The sensitivity was between 0.65 and 0.81 whilst specificity was between 0.85 and 0.97.

The low sensitivity (0.33) for rotator cuff pathology in the second step improved constantly after enhancing the pain map of this disorder. In the third step with testers, the sensitivity for rotator cuff was between 0.52 and 0.72. This means that the changes of pattern for rotator cuff pathology in the second phase were done correctly.

The current study divided the shoulder maps marked by the patients into two main groups; 1- pain localised to shoulder, 2- pain radiation beyond shoulder. Shoulder disorders were also classified as two main groups according to their pain patterns as above. The substantial agreement between map group and disease group in the second phase of the study means if anyone can estimate the group of the map using the shoulder pain mapping algorithm, they will match the group of the disease in over 80 % of cases. Estimating the disease group will narrow the number of the possibilities for the differential diagnoses and / or at least it will give an idea on what the diagnosis could be.

In summary, the reports of the results in the current study showed very good agreement within the disease groups using the mapping algorithm. Although the current study was a unique study with its content and there were no comparable studies in the literature, the sensitivities and specificities in testing shoulder pain maps were comparable to clinical tests for some shoulder disorders especially for instability, but also good for impingement syndrome, rotator cuff pathology and ACJ pathology. Due to the low number of patients, it was difficult to comment on the patients with GHJ arthritis and calcific tendinitis despite some good results. The algorithm has been enhanced with testers' advice.

In the third phase, inter-tester reliability test was used to observe a score of how much consensus or homogeneity there was for the algorithm of the shoulder pain maps. There was substantial agreement between the testers for grouped diseases and specific diseases. Inter-tester reliability helps to understand the consistency of the raters estimates (Cherry, 2013). The inter-tester reliability results in the third phase of the study with high agreement means there is consistent estimates of the algorithm and pain mapping system (Trochim, 2006). High agreement between the raters shows the clarity and the value of the algorithm and the established shoulder maps.

In the literature search, most of the relevant intertester reliability studies were on spinal pain and there was one study on shoulder pain. Two of them were to test the classification systems in spine and shoulder pains. Shoulder pain study showed the kappa coefficient was 0.664 (Carter, Hall, McIntosh, Murphy, MacDougall, & Boyle, 2010), and spine pain classification showed (kappa = 0.61) (Wilson, Hall, McIntosh, & Melles, 1999). The comments for both studies were that their kappa coefficients denoted good reproducibility.

Comparing to the above studies, the current study showed even higher Kappa (κ) coefficient (0.71).

The raters did not require long training or explanation, only a brief explanation. They commented that the algorithm and the colour-coded shoulder maps together were mostly self-explanatory. However, they also contributed to make further improvement in the algorithm, specifically on how to map the algorithm to assess the pins and needles or age distribution to ease the understanding of a third person.

Pain maps:

The pain mappings for shoulder disorders in the current study were extensively studied along the three phases. They were established, tested, improved and described well for each (seven) shoulder condition accordingly in this unique study. There were no previous studies on shoulder pain maps in the literature.

The shoulder pain maps reflected the patients' subjective experience, which can be considered as part of the history taking process. History taking was described as an art by (Rull & Draper, 2011). It is widely taught that diagnosis is revealed in the patient's history and sometimes it is all that is required to make the diagnosis. A well-designed visual pain map as in the current study's example, may guide patients to express their symptoms in a more descriptive way and in return, this may help doctors diagnosing shoulder problems. However, they would not be enough themselves for full diagnosis.

Pain maps have been found to be useful, for diagnostic, therapeutic, prognostic and research purposes (Palmer, 1949). Pain mapping has been previously described for neck, face, back and hip pain (Toomingas, 1999; Machacek & Friedrich, 2006; Turp, Kowalski, O`Leary, & Stohler, 1998). Pain maps were used for different applications.

There were various anatomical divisions of the body on the pain maps in the literature. Turp (1998) mentioned up to 50 different regions or anatomical sites were distinguished or identified on the body maps in previously published studies. In the current study, we used both front and back of the shoulder and arm with an extension towards neck (? Figure I). Although we described 14 cells on each side of the upper limb, for clinical and

diagnostic use, we divided the regions into only four in our descriptions; shoulder, down to elbow, below elbow area (forearm) and hand. The current study would not allow us to divide into more anatomical divisions as it would not be a practical approach for diagnostic purpose and would cause more conflict and confusion.

All the patients in the current study marked their pain on upper limb region. Although our custom-made shoulder map did not include the whole body mapping and excluded the pain related to neck as part of exclusion criteria, there were no comments from the patients regarding radiation of the pain to a different region outside upper limb. Whereas in the study by Turp et al. (1998), shoulder pain was described very often by the facial pain patients; the majority of the patients (69%) in the study reported pain outside the face and their results also show the neck, shoulders, and upper back to be the most frequently involved areas. The current study excluded patients with neck pain in the beginning, prior to patient recruitment to avoid mixed picture of diagnoses and to isolate the origin of the pain for closer match to shoulder diagnoses.

Shoulder pain maps may have a useful clinical role as an adjunct in the initial assessment of the shoulder problems as well as a research or follow-up tool. In any case, the definite diagnosis of shoulder problems could be very difficult despite the improvement in imaging technology. Symptoms and medical imaging may not correlate well (Carter et al., 2010) and investigations come with their cost. For example, ultrasound could be more cost-effective in a specialist hospital setting for identification of full-thickness tears comparing to MRI scan (Dinnes et al., 2003). Whilst cost effectiveness is important in investigating shoulder disorders, shoulder mapping can be an adjunct to the diagnosis. It can be particularly helpful in general practice. The practitioner, who may not have the expertise

in diagnosing the shoulder problems, may benefit more from the guidance of the pain mapping algorithm to prevent unnecessary investigations such as an MRI scan for ACJ pathology.

The use of pain maps in different settings may have different interpretations. According to the North American Spine (2013), when nerve roots are compressed, the pain signals they generate can be felt in different areas throughout the body. The pain mapping system can be useful to isolate the origin of the pain. However, the experience of pain is personal and it can be different for each individual (Brain Injury, 1998). The patient's description of pain is important and this may guide clinicians towards diagnosing the disorders.

There are some other studies which may give some ideas about how useful the pain maps could be. Machacek and Friedrich (2006) aimed to find out the reliability of several dermatomic maps (description of pain pattern within a nerve root lesion) according to the respective segmental area. They did not find a clear correlation between the pain projections of the lumbar spine. They concluded that the pain pattern of dermatomic maps is only of limited value for the definition of the affected segment. In a prospective study, however, Pang, Mok, Lin, Chang and Hwang (1998) examined and analysed 104 consecutive adult patients who underwent spinal pain mapping. They concluded that spinal pain mapping provided a useful functional approach to the diagnosis of low back pain with obscure aetiology in 87% of patients in their series. Likewise, when Wright (2000) conducted a study about referred craniofacial pain patterns in patients with temporomandibular disorder, he found that the pattern between referred pain source and site was consistent and predictable. As a result, all the studies above concluded that pain mapping is a useful and valuable tool for clinical assessment.

A different use of pain mapping was described for the human brain using functional magnetic resonance imaging (fMRI). It is suggested that pain maps could be used to determine the kind of pain which is felt by patients who are unable to articulate it (RT, 2013). The pain maps could be used to provide markers for the location in the human brain and this would enable clinicians to understand how patients' brain reorganise following chronic pain (Medicalxpress, 2012). In appropriate setting, questionnaires may give very useful information about disorders. For example, in the assessment of neck and upper extremity disorders, Ohlsson et al., (1994) concluded that the questionnaire approach gives a fairly good picture of the neck / upper extremity status of a working female population. Most subjects with findings on clinical examination of shoulders reported symptoms on the questionnaire (sensitivity 80%).

Visual analogue scales (VAS) was a part of the mapping questionnaire in the current study and it was used to compare the severity of pain. The VAS scales create ratio level data that is more easily comparable (Disorbio & Bruns, 1999). Although its value is limited by the previous lack of a standardized method for performing such a rating, or any scientifically-based norms for making comparisons, the 0-10 pain scale had benefit from an extremely broad acceptance in the field (Disorbio & Bruns, 1999).

Toomingas (1999) looked at subjects with chronic and severe pain and correlated this with pain drawings. Patients with more chronic or severe pain symptoms had pain drawings that occupied a larger area. As suggested above, our study also showed the same correlation existence between the severity of pain and area of radiation. The most severe pain was in GHJ arthritis patients, who marked a mean of 12.5 out of 28 cells on the pain map, which was the greatest number amongst all groups. Moreover, the ACJ pathology

was the most pinpoint described one and the average of area marked by the patients was only 3.9 with the least severe pain of 6.43 on VAS.

In the current study, pain patterns were constant and very predictable for the disease groups (1 and 2) as well as showing good results for individual shoulder diseases especially instability, ACJ pathology and impingement. The substantial agreement between the raters shows the reliability of the algorithm and colour-coded shoulder pain maps.

This is the first study to describe the usage of pain maps in the shoulder and a unique study with its three phases. There was no previous study on any of the three phases of shoulder pain maps. In the literature, the studies on pain maps of face, hip or back were with evidence level of four but not blinded and no review studies. In the current study, a definitive pattern of pain distribution and specific types of pain in common shoulder pathologies has been demonstrated. One advantage of the shoulder mapping could be that the patients might complete the questionnaire before the clinic visit and the clinician would have an idea about the problem before seeing the patient. It is simpler to understand and easier to complete as a visual questionnaire and possibly more advantageous over the written one.

Limitation:

In an inter-observer agreement study for classification of shoulder disorders, de Winter et al. (1999) concluded that differentiation between the various categories of shoulder disorders is complicated. The number of the disorders (seven) and different types of pain (four) meant too many variables and made the study more difficult for statistical approach.

This led to 'hand analysis' of disorder patterns directly from the excel sheet. Additionally, the small number of patients with some diagnosis such as GHJ arthritis and calcific tendinitis was another limitation of the study related to analysis of those disorders. Therefore, it caused a problem with testing of the pain mapping, although this could be a reflection of frequency of the disorders in the population.

Larger sample size would give further accuracy in shoulder pain mapping especially for GHJ arthritis and calcific tendinitis. However, the study was blinded study in terms of diagnoses. The number of the groups and the number of the patients in each group were not known before obtaining final diagnoses for the included patients. Therefore, sample size was not calculated prior to the phase one and two. In addition to six diagnoses in the first phase of the study, another diagnosis was added (frozen shoulder) in the second phase after obtaining the final diagnoses. But post-hoc calculations were performed for the results with statistically significant findings.

Further exclusion of the patients in the second phase was a major limitation to the study and this possibly contributed to the low number the above diagnoses. Despite the initial inclusion and exclusion criteria, many patients (128) were excluded from the second phase of the study at a later stage. Some of these exclusions were secondary to patients misinterpreting how to mark the custom-made shoulder map form despite verbal explanations. Whenever possible, we included the maps to the study but we did not want to cause bias. Therefore, we excluded the unacceptable maps. The other major reason for exclusion despite initial inclusion was inadequate clinical data. This probably was due to patients been lost to follow up or drop out. Drop outs maybe led to inadequate clinical information of those patients for accurate diagnoses. Therefore, they were excluded from the study. In turn, this decreased the number of the patients affected the sample size.

Testing the color-coded shoulder pain maps showed better results for ACJ and instability in the second phase, however, the results showed that it was not very good for rotator cuff pathology. This was probably due to the similar pain patterns of two or three disorders. After the necessary changes to its pattern at the end of phase two, the third phase showed better results. A similarity in some pain patterns of shoulder disorders was another limitation to define precise pain maps for shoulder disorders. This led to classify the pain patterns into groups and subgroups according to their similarities.

Although the study was conducted in different hospitals and it was tested by three different raters in addition to the researcher, multi-centre studies might lead to better generalisation.

The study aimed to perform a quantitative study about shoulder pain mapping. An addition of qualitative assessment to the study about pain perception of the patients would increase the strength of the study and potentially give a better picture of pain patterns. However, this will probably mean further extension of the study. It could be beyond the scope the current study.

The other limitation was misinterpretation of the pain maps by the patients although the map itself was very explanatory and simple, and the patients were explained the map before they marked them, there were still some mismarked maps.

No age limit was planned and it showed massive age ranges. The clinic included all the adult population with shoulder problems. For example GHJ arthritis, ACL pathologies are commoner in older patients, whereas instability is commoner is younger patients. If we limit the age, we would miss a variety of the diseases and the number of the patients included in the study would be less.

Conclusion:

The authors of the study advocate the use of pain maps as a diagnostic tool in shoulder clinics and this can also be very helpful in the primary care setting. Although shoulder mapping itself is insufficient for precise clinical diagnosis, it can be valuable as an adjunct to the other assessments and play an important role in shoulder disorders

This is a unique study with its content and conclusion. There was no other previous study to compare the results with. We believe that this study will fill the gap for the patient assessment from the patient subjective experience to shoulder pathology.

This is the first study to describe the use of pain maps in the shoulder. A definitive pattern of pain distribution and specific types of pain in common shoulder pathologies has been demonstrated. The study may further suggest that accurate history from patients about their symptoms and pain distribution may give important clues about their diagnosis and may potentially prevent unnecessary investigations.

Further Use:

It can be remotely used as online - computer tools for the patients self assessments, or as pre-attendance to clinics, which may help the clinician have an idea about the shoulder problem when the maps are utilised with the algorithm.

The use of pain maps could be further expanded as a diagnostic aid in patients with combined neck and shoulder pathology and multiple shoulder pathologies. It may lead to further studies for example, combination of shoulder mapping with psychological assessment tools to convey the recognition of an additional psychological approach.

References:

Altman, D. G., Bland, J. M. (1994a). General practice Statistics Notes: Diagnostic tests 2: predictive values. *British Medical Journal*, 309:102.1
Retrieved 27 August, 2013, from
http://www.bmj.com/content/309/6947/102.1

Altman, D. G., Bland, J. M. (1994b). Diagnostic tests. 1: Sensitivity and specificity. *British Medical Journal,* 308(6943),1552.

Anthony, D. (1999) *Understanding advanced statistics*. London: Churchill Livingstone.

Apley, A. G., Solomon, L. (1993). *Diagnosis in Orthopaedics.* In: Apley`s System of Orthopaedics and Fracture (pp.3-4). (7th ed.) Oxford: Butterworth Heinemann.

Baird, J. C. (2009). Three-Dimensional Pain Mapping. (2009) Retrieved 27 July, 2013, from http://www.sbir.gov/sbirsearch/detail/283557

Bayam, L., Ahmad, M. A., Naqui, S. Z., Chouhan, A., & Funk, L. (2011). Pain mapping for common shoulder disorders. *American Journal of Orthopaedics*, 40(7), 353-358.

Bjelle, A. (1989). Epidemiology of shoulder problems. *Balliere's Clinical Rheumatology*, 3, 437-451.

Brain Injury Resource Center. (1998). Pain Map. Retrieved 15 November, 2013, from http://www.headinjury.com

Brue, S., Valentin, A., Forssblad, M., Werner, S., Mikkelsan, C., & Cerulli, G. (2007) Idiopathic adhesive capsulitis of the shoulder: a review. *Knee Surgery, Sports Traumatology, Arthroscopy*, 15(8), 1048-1054.

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), 453-60.

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, 44-47. doi:10.1136/ard.59.1.44

Carter, T., Hall, H., McIntosh, G., Murphy, J., MacDougall, J., & Boyle, C. (2010). Intertester reliability of a classification system for shoulder pain. Physiotherapy, 98(1), 40-46.

Cherry, K. (2013) What Is Reliability? Retrieved 18 September, 2013, from http://psychology.about.com/od/researchmethods/f/reliabilitydef.htm

Clinical Trials. U.S. National Institutes of Health (2012). Glossary of Common Site Terms. Retrieved 08 August, 2013, from http://clinicaltrials.gov/ct2/about-studies/glossary#masking-or-blinding

De Winter, A. F., Jans, M. P., M. P., Scholten, R. J. P. M., Devillé, W., Schaardenburg, D. V., & Bouter, L. M. (1999). Diagnostic classification of shoulder disorders: interobserver agreement and determinants of disagreement. *Annals of the Rheumatic Diseases*, 58, 272–277.

Di Massa, A., Avella, R., & Gentili, C. (1996). Respiratory dysfunction related to diaphragmatic shoulder pain after abdominal and pelvic laparoscopy. *Minerva Anestesiologica*, 62(5), 171-176.

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 Journal*, 7(29), iii, 1-166.

Disorbio, J. M., & Bruns, D. (1999). *Diagnosis Specific Norms for the 0-10 Pain Scale*. Presented at World Congress on Pain. Vienna, Austria 1999.

Douglas, J. (2006). Pain mapping software helps patients describe pain to doctors with 3-D graphics. Retrieved 18 September, 2012, from http://www.naturalnews.com/021034_medical_technology_pain_management.html

Dutton, M. (2008). *Orthopaedic Examination, Evaluation, and Intervention* (2nd Ed.). New York: McGraw-Hill.

Faragher, B. (2005). *Essential Statistics for Medical Examinations* (2nd Ed.). Cheshire: Pastest Limited.

Fodor, D., Poanta, L., Felea, I., Rednic, S., & Bolosiu H. (2009). Shoulder impingement syndrome: correlations between clinical tests and ultrasonographic findings. *Ortopedia Traumatologia Rehabilitacja*, 11(2), 120-126.

Fraser-Moodie, J. A., Shortt, N. L., & Robinson C M. (2008). Injuries to the acromioclavicular joint. *Journal of Bone and Joint Surgery (British volume)*, 90(6), 697-708.

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.

Guanche, C. A., & Jones, D. C. (2003) Clinical testing for tears of the glenoid labrum. *Arthroscopy*, 19(5), 517-523.

Harris, M. & Taylor, G. (2008). *Medical Statistics, Made easy*, (2nd ed.), (pp. 67-68). Oxfordshire, UK: Scion Publishing Limited.

Hassett, G., & Barnsley, L. (2001). Pain referral from the sternoclavicular joint: a study in normal volunteers. *Rheumatology (Oxford*), 40(8), 859-62.

Health Research Authority (2013) Requirements for ethical review under legislation. National Research Ethics Service (NRES). Retrieved 07 August, 2012, from

http://www.nres.nhs.uk/applications/approval-requirements/ethical-review-requirements/requirements-for-ethical-review-under-legislation/

Hegedus, E. J., Goode, A., Campbell, S., Morin, A., Tamaddoni, M., Moorman, C. T. 3rd, & Cook, C. (2008). Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *British Journal of Sports Medicine.* 42(2), 80-92.

Hróbjartsson, A., Thomsen, A. S. S., Emanuelsson, F., Tendal, B., Hilden, J., Boutron, I., Ravaud, P., & Brorson, S. (2013). Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *Canadian Medical Association Journal*, 185, E201-E211. doi:10.1503/cmaj.120744

Hurt, G., & Baker, C. L. (2003).). Calcific tendinitis of the shoulder. *Orthopaedic Clinics of North America,* 34(4), 567-575.

Inman, V. T., & Saunders, J. B. (1944). Referred pain from skeletal structures. *Journal of Nervous and Mental Disease*, 99, 660-667.

International Association for the Study of Pain (IASP) (2012). Taxonomy. Retrieved 27 March, 2014, from https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1698

International Committee of Medical Journal Editors (1993) 'Conflict of interest' *Annals of Internal Medicine*, 118(8), 646-647.

Jensen, M. P., Karoly, P. (1991). Pain-specific beliefs, perceived symptom severity, and adjustment to chronic pain. *Clinical Journal of Pain*, 8, 123-130.

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.

Kovach, S. G., & Huslig, E, L. (1984). Shoulder pain and Pancoast tumor: a diagnostic dilemma. Journal of Manipulative and Physiological Therapeutics, 7(1), 25-31.

Larson, H. M., O`Connor, F. G., & Nirschl, R. P. (1996). Shoulder pain: the role of diagnostic injections. *American Family Physician*, 53(5), 1637-1647.

Limb, D., & Hay, S. M. (2007). The evidence for Orthopaedic Surgery. Shrewsbury, UK: TFM Publishing Limited.

Lowry, R. (2013). Sensitivity and specificity. Retrieved 08 August, 2013, from http://vassarstats.net/clin1.html

Lundberg, B. (1969). The frozen shoulder. *Acta Orthopaedica Scandinavia*, Suppl 119.

Machacek, P., & Friedrich, M. (2006). Pain patterns of disc lesions of the lumbar spine. *Z Orthop Ihre Grenzgeb.*, 144(1), 19-26.

Mall, N. A., Foley, E., Chalmers, P. N., Cole, B. J., Romeo, A. A., & Bach, B. R. (2013) Degenerative Joint Disease of the Acromioclavicular Joint: A Review. *American Journal of Sports Medicine*, 41(11), 2684-2692.

Medicalxpress. (2012). Where does it hurt? Pain map discovered in the human brain. Retrieved 06 July, 2013, from http://medicalxpress.com/news/2012-11-pain-human-brain.html

Meislin, R. J., Sperling, J. W., & Stitik, T. P. (2005). Persistent shoulder pain: epidemiology, pathophysiology, and diagnosis. *American Journal of Orthopedics*, 34(12), Suppl(5-9), 1078-4519.

Methodology tables (2002) Retrieved 20 April, 2014, from http://www.health.ny.gov/community/infants_children/early_intervention/disorders/appendix_a.htm

Microsoft Corporation (2010) 'Microsoft Office Word' About spelling, grammar, and the thesaurus. Retrieved 06 April, 2013, from http://office.microsoft.com/en-us/word/HP052366411033.aspx

Milewski, M. D., Hart, J. A., & Miller, M. D. (2012). Sport Medicine In: Miller, M. D., Thompson, S. R., & Hart, J. A. (Eds) *Review of Orthopaedics.* (6th ed.) (pp. 319-334). Philedelphia, US: Saunders, Elsevier.

Mitchell, C., Adebajo, A., Hay, E., & Carr, A. (2005). Shoulder pain: diagnosis and management in primary care. Clinical review. *British Medical Journal*, 331, 1124-1128.

National Cancer Institute. (2012). NCI Dictionary of Cancer Terms. Retrieved 08 August, 2013, from
http://www.cancer.gov/dictionary?cdrid=285982

Neviaser, R. J., & Neviaser, T. J. (1987) The frozen shoulder. Diagnosis and management. *Clinical Orthopaedics and Related Research* , 223, 59-64.

Newman, T. B., Browner, W. S. & Hulley, S. B. (2001) `Enhancing causal inference in onservational studies` in: Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D., Hearst, N. & Newman, T. B. *Designing Clinical Research,* (2nd ed). Philadelphia: Lippincott Williams&Wilkins. pp.125-138.

NHS Evidence (provided by NICE) (2009) 'NLH Resources' Health Information Resources. Retrieved 06 April, 2013, from
http://www.library.nhs.uk/help/resource

North American Spine. (2013). Pain Mapping. Retrieved 27 July, 2013, from
http://northamericanspine.com/spine-procedure/diagnosis/pain-mapping/

Ohlsson, K., Attewell, R. G., Johnsson, B., Ahlm, A., & Skerfving, S. (1994). An assessment of neck and upper extremity disorders by questionnaire and clinical examination. *Ergonomic,* 37(5), 891-897.

Pain management (2014). Retrieved 27 March, 2014, from http://paincenter.stanford.edu/patient_care/pain.html

Pain signature (2013). Scientists map aching and its intensity in brain. Retrieved 06 July, 2013, from http://rt.com/news/pain-signature-brain-scan-675/

Palmer, H. (1949). Pain charts. A description of a technique whereby functional pain may be distinguished from organic pain. *New Zealand Medical Journal,* 48, 187-213.

Pang, W. W., Mok, M. S., Lin, M. L., Chang, D. P., & Hwang, M. H. (1998). Application of spinal pain mapping in the diagnosis of low back pain--analysis of 104 cases. *Acta anaesthesiologica Sinica,* 36(2),71-74.

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 volume),* 87(7), 1446-1456.

Polgar, S. & Thomas, S. A. (1991) *Introduction to Research in the Health Science.* (2nd ed.). Melbourne: Churchill Livingstone.

ProQuest (2010) 'Medline' Key facts: Medline. Retrieved 06 April, 2013, from http://www.proquest.com/en-US/catalogs/databases/detail/medline_ft.shtml

Rankine, J.J., Fortune, D. G., Hutchinson, C. E., Hughes, D. G., & Main, C. J. (1998). Pain drawings in the assessment of nerve root compression: a comparative study with lumbar spine magnetic resonance imaging. *Spine,* 23, 1668-1676.

Richards, R. R., Kai-Nan An, C., Bigliani, L. U., Friedman, R. J., Gartman, G. M., Gristina, A. G., Iannotti, J. P., Mow, V. C., Sidles, J. A., & Zuckeman, J. D. (1994). A standardized method for the assessment of shoulder function. *Journal of Shoulder and Elbow Surgery,* 3(6), 347–352.

Rodowsky, M. W. and Bigliani, L.U. (1997). The Shoulder and the Shoulder Girdle. In: Dee, R., Hurst, L. C., Gruber, M. A., & Kottmeier, S. A (Eds). *Principles of Orthopaedic Practice*. (2nd ed.). (pp. 1063). New York: McGraw–Hill.

Rull, G., & Colin, T. (2013). Acromioclavicular Joint Problems. Retrieved 03 July, 2013, from
http://www.patient.co.uk/doctor/acromioclavicular-joint-problems

Rull, G., & Draper, R. (2011) The art of history taking Retrieved from http://www.patient.co.uk/doctor/history-taking

Schwarzkopf, R., Oron, A., & Loebenberg, M. (2008). Shoulder pain: assessment, diagnosis and treatment of common problems. *Harefuah,* 147(1), 71-76, 93.

Sizer-Phillip, S., Phelps, V., & Gilbert K. (2003). Diagnosis and management of the painful shoulder. Part 2: examination, interpretation, and management. *Pain Practice,* 3(2), 152-185.

StatsDirect Limited. (2013) Prospective vs. Retrospective Studies. Retrieved 26 July, 2013 from

http://www.statsdirect.com/webhelp/#basics/prospective.htm

Toomingas, A. (1999). Characteristics of pain drawings in the neck-shoulder region among the working population. *International Archives of Occupational and Environmental Health*, 72(2), 98-106.

Turp, J. C., Kowalski, C. J., O`Leary, N., & Stohler, C. S. (1998). Pain maps from facial pain patients indicate a broad pain geography. *Journal of Dental Research*, 77(6), 1465-1473.

Trochim, W. M. K. (2006). Types of Reliability. Retrieved 08 August, 2013, from http://www.socialresearchmethods.net/kb/reltypes.php

Tzannes, A., & Murrell, G. A. (2002). Clinical examination of the unstable shoulder. *Sports Medicine,* 32(7), 447-457.

Ure, B. M., Tiling, T., Kirchner, R., & Rixen, D. (1993). Reliability of clinical examination of the shoulder in comparison with arthroscopy. *Unfallchirurg*, 96(7), 382-6.

Urwin, M., Symmons, D., Allison, T., Brammah, T., Busby, H., Roxby, M., Simmon, A., 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 Woude, H. J., & Vanhoenacker, F. M. (2007). MR arthrography in glenohumeral instability. *Journal Belge de Radiologie - Belgisch Tijdschrift voor Radiologi*, 90(5), 377-383.

Wilson, L., Hall, H., McIntosh, G., Melles, T. (1999) Intertester reliability of a low back pain classification system. Spine (Phila Pa 1976), 24(3), 248-254.

Woodward, T. W., & Best, T. M. (2000) The Painful Shoulder: Part II. Acute and Chronic Disorders. *American Family Physician*, 61(11), 3291-3300.

Woodward., A. H. (2013). Calcifying Tendonitis. Retrieved 15 September, 2013, from http://emedicine.medscape.com/article/1267908-overview#a0112

Wright, E. F. (2000). Referred craniofacial pain patterns in patients with temporomandibular disorder. *Journal of the American Dental Association,*131(9),1307-15.

Appendices

Wrightington, Wigan and Leigh NHS
NHS Foundation Trust

Research and Development Department
Wrightington Hospital
Hall Lane
Appley Bridge
Wigan
WN6 9EP

Mrs Sandra Latham – Research and Development Co-ordinator
Mrs Christine Birchall – Research and Development Manager
Prof Nirmal Kumar, MPhil, DLO, FRCS (ORL-HNS) – Clinical Director of R & D

Tel: 01257 256465
Fax: 01257 256398

Sandra.Latham@wwl.nhs.uk

Mr L Bayam and Prof L Funk
Upper Limb Research
Wrightington Hospital
Wigan

20th June 2011

Dear Mr Bayam and Prof Funk

RE: Shoulder Pain Mapping – Questionnaire Study

This letter is confirmation that you have Research and Development approval to conduct the above titled research study within Wrightington, Wigan and Leigh NHS Foundation Trust.

Once again I would like to draw your attention to the following Trust policies:

Health and Safety at Work Act, Data Protection Act 1998, Human Tissue Act 2004: Good Clinical Practice.

It is a requirement that the Research and Development Department are informed of:

- Any Amendments arising from the research study.
- Any Serious Adverse Events (SAE's) that might arise during the course of the study.
- Any presentations or publications arising from the study.

Yours sincerely

Prof Nirmal Kumar
Clinical Director of Research and Development
Consultant Otolaryngologist – Head and Neck Surgeon

shoulder*				
AND				
Disease*	OR	Disorder*		
AND				
Pain*				
AND				
Map*	OR	Draw*	OR	Pattern*

Table -1: Search Terms

Literature Search Results of Databases

1-) The Cochrane Library: 09/12/2013 via the Wiley Online Library (repeat search)

http://onlinelibrary.wiley.com/cochranelibrary/search

Cochrane Database of Systematic Reviews : Issue 12 of 12, December 2013

There are 58 results from 8219 records for your search on 'shoulder pain' in Title, Abstract, Keywords in Cochrane Reviews'

Cochrane Central Register of Controlled Trials : Issue 11 of 12, November 2013

There are 34 results from 717246 records for your search on 'shoulder pain pattern' in Title, Abstract, Keywords in Trials'

2- MEDLINE (from 1950 to 04/12/2013) via NHS Evidence (www.library.nhs.uk)

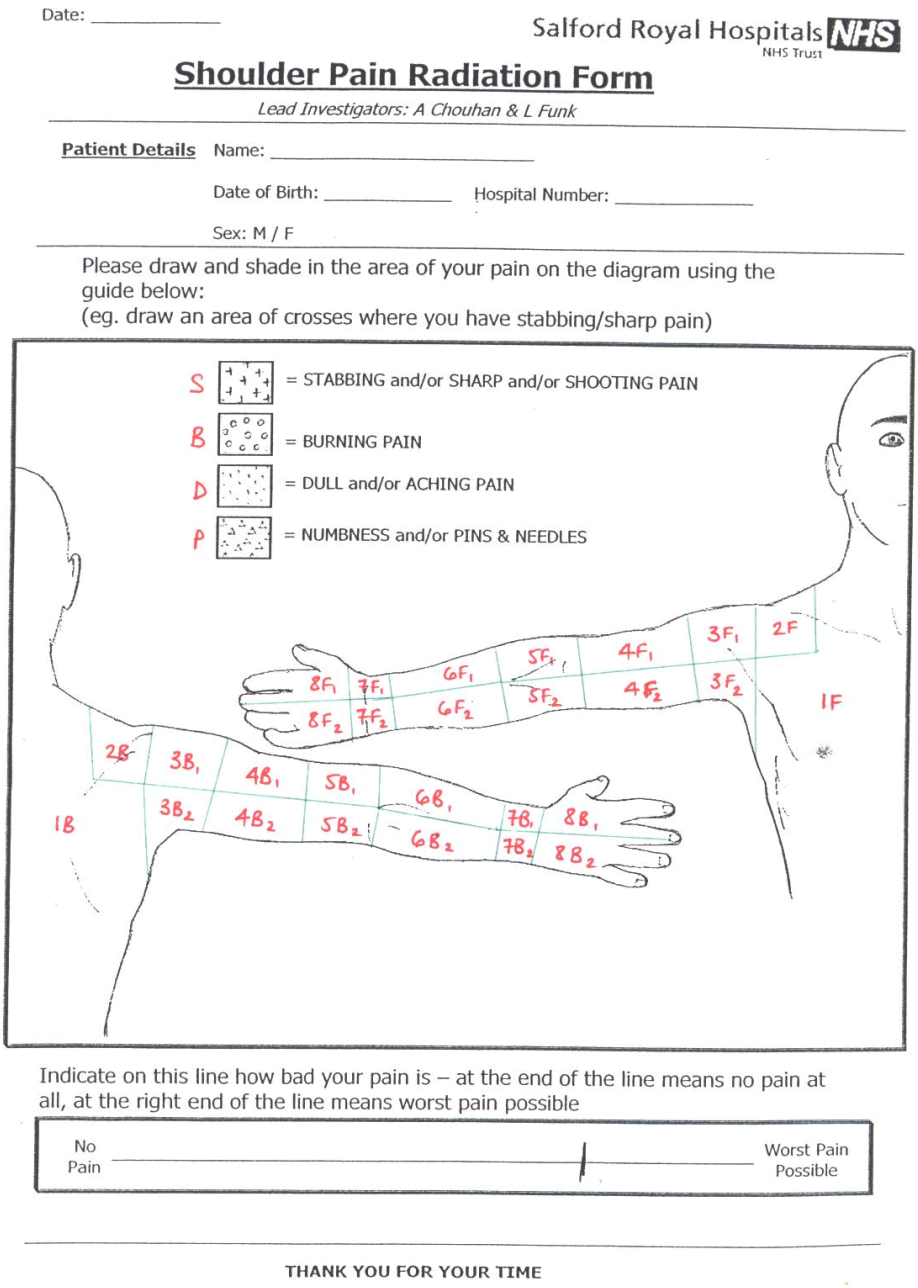
No.		Search term	Hits
1	MEDLINE	exp SHOULDER/	8872
2	MEDLINE	shoulder.ti,ab	41588
3	MEDLINE	1 OR 2	44979
4	MEDLINE	exp PAIN/ OR exp SHOULDER PAIN/	310095
5	MEDLINE	pain.ti,ab	408320
6	MEDLINE	4 OR 5	547510
7	MEDLINE	disorder*.ti,ab	742827
8	MEDLINE	exp DISEASE/	117746
9	MEDLINE	disease*.ti,ab	2570479
10	MEDLINE	7 OR 8 OR 9	3184337
11	MEDLINE	map*.ti,ab	335803
12	MEDLINE	pattern*.ti,ab	936017
13	MEDLINE	draw*.ti,ab	136400
14	MEDLINE	11 OR 12 OR 13	1362958
15	MEDLINE	3 AND 10	6112
16	MEDLINE	6 AND 14	23918
17	MEDLINE	15 AND 16	184
18	MEDLINE	Duplicate filtered: [15 AND 16]	184 171 Unique results 13 Duplicate results

3- EMBASE (from 1980 to 04/12/2013) via NHS Evidence (www.library.nhs.uk)

No.		Search term	Hits
1	EMBASE	exp SHOULDER/	21574
2	EMBASE	shoulder.ti,ab	48471
3	EMBASE	1 OR 2	54612
4	EMBASE	disorder*.ti,ab	897555
5	EMBASE	disease*.ti,ab	3064527
6	EMBASE	4 OR 5	3695788
7	EMBASE	exp LIMB PAIN/ OR exp PAIN/ OR exp PAIN ASSESSMENT/ OR exp SHOULDER PAIN/	796051
8	EMBASE	pain*.ti,ab	587424
9	EMBASE	7 OR 8	1007646
10	EMBASE	map*.ti,ab	350312
11	EMBASE	pattern*.ti,ab	990572
12	EMBASE	exp DRAWING/	2091
13	EMBASE	draw*.ti,ab	158650
14	EMBASE	10 OR 11 OR 12 OR 13	1451949
15	EMBASE	3 AND 6	7841
16	EMBASE	9 AND 14	42440
17	EMBASE	15 AND 16	276
18	EMBASE	Duplicate filtered: [15 AND 16]	276 265 Unique results 11 Duplicate results

CINAHL (from 1981 to 04/12/2013) via NHS Evidence (www.library.nhs.uk)

No.	Search term		Hits
1	CINAHL	exp SHOULDER/	3130
2	CINAHL	shoulder.ti,ab	9223
3	CINAHL	exp DISEASE/	76292
4	CINAHL	disease*.ti,ab	200778
5	CINAHL	disorder*.ti,ab	77565
6	CINAHL	1 OR 2	10173
7	CINAHL	3 OR 4 OR 5	324229
8	CINAHL	exp PAIN/	90189
9	CINAHL	pain*.ti,ab	94255
10	CINAHL	8 OR 9	129047
11	CINAHL	exp MAPS/	1447
12	CINAHL	map*.ti,ab	11228
13	CINAHL	pattern*.ti,ab	53773
14	CINAHL	draw*.ti,ab	22274
15	CINAHL	11 OR 12 OR 13 OR 14	86083
16	CINAHL	6 AND 7	1601
17	CINAHL	10 AND 15	5011
18	CINAHL	16 AND 17	49
19	CINAHL	Duplicate filtered: [16 AND 17]	49 49 Unique results 0 Duplicate results



Date: _____

Salford Royal Hospitals NHS
NHS Trust

Shoulder Pain Radiation Form

Lead Investigators: A Chouhan & L Funk

Patient Details Name: _____

Date of Birth: _____ Hospital Number: _____

Sex: M / F

Please draw and shade in the area of your pain on the diagram using the guide below:
(eg. draw an area of crosses where you have stabbing/sharp pain)

Indicate on this line how bad your pain is – at the end of the line means no pain at all, at the right end of the line means worst pain possible

No Pain _____ Worst Pain Possible

THANK YOU FOR YOUR TIME

Figure I: The sample map to be marked by the patients (excluding the red and green marks: they were only used on comparison sheet.)

The Colour-coded Pain Maps for Common Shoulder Disorders

Colour coding of pain

(Bayam et al, 2011)

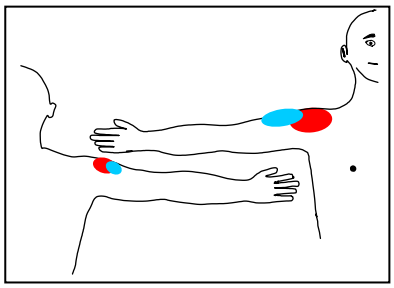


Figure 1: ACJ pathology

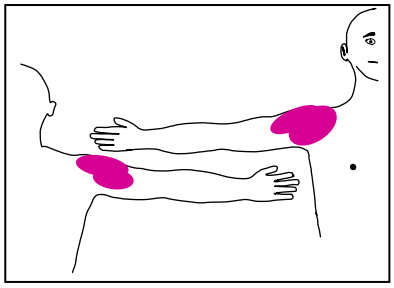


Figure 2: Instability

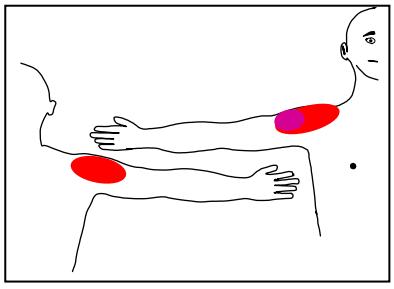
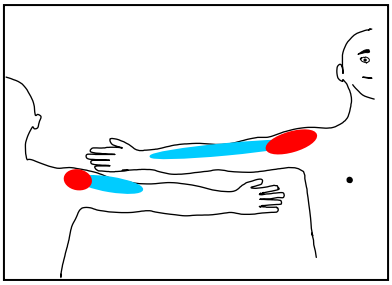
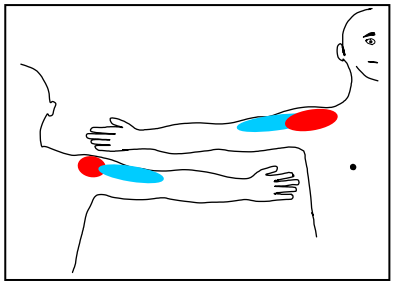


Figure 3: Calcific tendinitis



4a



4b

Figure 4a and 4b: Rotator cuff tear (the upper one- 4a is from the first phase
The lower one- 4b is from the second phase)

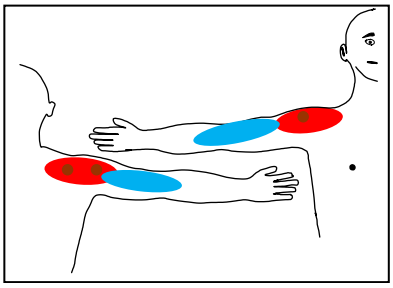


Figure 5: Frozen Shoulder

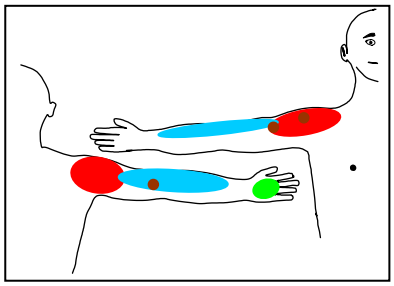


Figure 6: Impingement

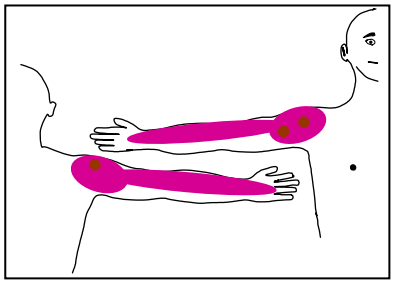
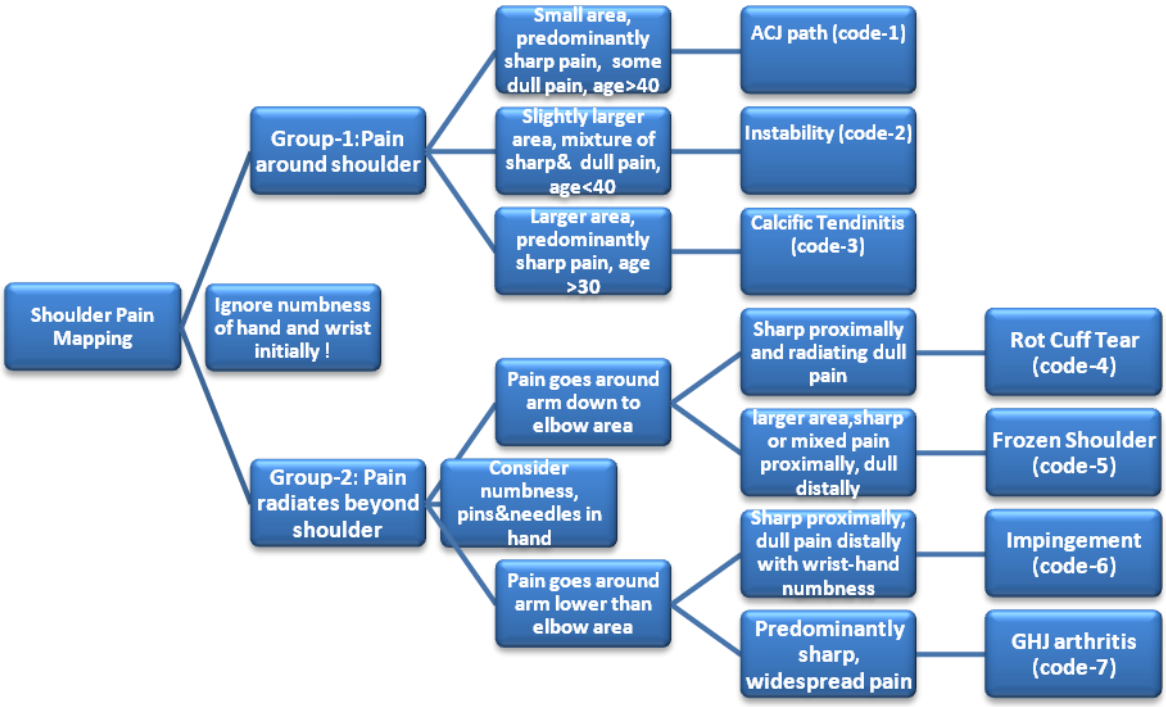


Figure 7: GHJ Artritis

.

Shoulder Pain Mapping Algorithm



The extension of the marked area of pain from less to more (the numbers shows the disorder codes):

1 < 2 < 3 < 4 < 5 < 6 < 7

How to Read the Map:

1- Initially, the shoulder disorders are divided into two main group from map point of view:

First group: The ones with the pain **around only shoulder** (ACJ path, Instability, Calc Tendinitis), limited radiation of pain beyond shoulder

Second group: the ones with the pain **radiates beyond the shoulder** (Rotator cuff tear, Frozen Shoulder, Impingement, GHJ arthritis).

2- First step is to look if marking finishes above, below or around elbow **– ignore the numbness around hand and wrist**

Around Shoulder, above elbow: ACJ pathology(Code-1), Instability (Code-2), Calcific tendinitis (Code-3)

The extension of the marked area of pain from less to more (the numbers shows the disorder codes):

1 < 2 < 3

3- **A**- Small area of predominantly sharp pain, maybe some dull pain→ ACJ (Code-1)

4- **B**- Larger area of mixture of sharp and dull pain → instability (Code-2)

5- If any indecisiveness between instability and ACJ pathology, if age <40 most probably instability and >40 most probably ACJ pathology

6- **C**- Larger area, but predominantly sharp pain → Calcific tendinitis(Code-3)

7- The second group is subdivided into:

I- **Around arm down to elbow level**: **D**-rotator cuff tear (Code-4), **E**-frozen shoulder (code-5)

II- **Below elbow level**: **F**- Impingement (Code-6), **G**-GHJ arthritis (Code-7)

8- **D**- Sharp proximally and radiating as dull pain distally→ rotator cuff (Code-4)

9- **E**- Characterized with sharp or mixed pattern proximally and dull pain more distally → Frozen shoulder (Code-5)

10- The extension of the marked area of pain from less to more (the numbers shows the disorder codes):

4 < 5

11- The disorders that the marking goes beyond elbow

12- **F**- Sharp pain proximally and dull pain distally with wrist-hand numbness → Impingement (Code-6)

13- **G**- Predominantly sharp, widespread pain → GHJ arthritis (Code- 7)

14- The extension of the marked area of pain from less to more (the numbers shows the disorder codes):

6 < 7

15- The extension of the marked area of pain from less to more (the numbers shows the disorder codes):

1 < 2 < 3 < 4 < 5 < 6 < 7
Or
A < B < C < D < E < F < G

Full Statistics reports of the first phase

```
ONEWAY Painseverity Areano BY Diagnosis
  /STATISTICS DESCRIPTIVES HOMOGENEITY
  /PLOT MEANS
  /MISSING ANALYSIS
  /POSTHOC=TUKEY ALPHA(0.05).
```

Oneway

```
[DataSet1] G:\wrightington start\shoulder mapping\shoulder SPSS.sav
```

Descriptives

						95% Confidence Interval for Mean		
		N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum
Severity of Pain	impingement	27	7.5185	1.80534	.34744	6.8043	8.2327	3.00
	rotator cuff tear	22	7.0455	1.46311	.31194	6.3967	7.6942	5.00
	G. Arthritis	6	7.8333	1.94079	.79232	5.7966	9.8701	5.00
	Instability	18	6.7222	1.80866	.42630	5.8228	7.6216	3.00
	ACJ pathology	14	6.4286	1.74154	.46545	5.4230	7.4341	2.00
	Calcific tendinitis	6	7.5000	2.42899	.99163	4.9509	10.0491	3.00
	Total	93	7.1075	1.77826	.18440	6.7413	7.4738	2.00
Number of the areas	impingement	28	7.2143	3.83316	.72440	5.7279	8.7006	1.00
	rotator cuff tear	22	5.5909	3.37581	.71973	4.0942	7.0877	1.00

	G. Arthritis	6	12.5000	8.52643	3.48090	3.5521	21.4479	4.00
	Instability	17	6.8824	6.17335	1.49726	3.7083	10.0564	1.00
	ACJ pathology	14	3.8571	1.91581	.51202	2.7510	4.9633	2.00
	Calcific tendinitis	6	5.5000	4.03733	1.64823	1.2631	9.7369	2.00
	Total	93	6.4946	4.76998	.49462	5.5123	7.4770	1.00

Multiple Comparisons

Tukey HSD

Dependent Variable	(I) shoulder diseases	(J) shoulder diseases	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Severity of Pain	impingement	rotator cuff tear	.47306	.50884	.938	-1.0098	1.9559
		G. Arthritis	-.31481	.79960	.999	-2.6450	2.0154
		Instability	.79630	.53909	.680	-.7747	2.3673
		ACJ pathology	1.08995	.58347	.429	-.6104	2.7903
		Calcific tendinitis	.01852	.79960	1.000	-2.3117	2.3487
	rotator cuff tear	impingement	-.47306	.50884	.938	-1.9559	1.0098
		G. Arthritis	-.78788	.81596	.928	-3.1658	1.5900
		Instability	.32323	.56306	.992	-1.3177	1.9641
		ACJ pathology	.61688	.60569	.911	-1.1482	2.3820
		Calcific tendinitis	-.45455	.81596	.993	-2.8324	1.9233

	G. Arthritis	impingement	.31481	.79960	.999	-2.0154	2.6450
		rotator cuff tear	.78788	.81596	.928	-1.5900	3.1658
		Instability	1.11111	.83516	.767	-1.3227	3.5449
		ACJ pathology	1.40476	.86447	.584	-1.1145	3.9240
		Calcific tendinitis	.33333	1.02286	.999	-2.6475	3.3142
	Instability	impingement	-.79630	.53909	.680	-2.3673	.7747
		rotator cuff tear	-.32323	.56306	.992	-1.9641	1.3177
		G. Arthritis	-1.11111	.83516	.767	-3.5449	1.3227
		ACJ pathology	.29365	.63132	.997	-1.5462	2.1335
		Calcific tendinitis	-.77778	.83516	.937	-3.2116	1.6561
	ACJ pathology	impingement	-1.08995	.58347	.429	-2.7903	.6104
		rotator cuff tear	-.61688	.60569	.911	-2.3820	1.1482
		G. Arthritis	-1.40476	.86447	.584	-3.9240	1.1145
		Instability	-.29365	.63132	.997	-2.1335	1.5462
		Calcific tendinitis	-1.07143	.86447	.816	-3.5907	1.4478
	Calcific tendinitis	impingement	-.01852	.79960	1.000	-2.3487	2.3117
		rotator cuff tear	.45455	.81596	.993	-1.9233	2.8324
		G. Arthritis	-.33333	1.02286	.999	-3.3142	2.6475
		Instability	.77778	.83516	.937	-1.6561	3.2116

		ACJ pathology	1.07143	.86447	.816	-1.4478	3.5907
Number of the areas	impingement	rotator cuff tear	1.62338	1.27357	.798	-2.0881	5.3348
		G. Arthritis	-5.28571	2.01101	.101	-11.1462	.5748
		Instability	.33193	1.37446	1.000	-3.6735	4.3374
		ACJ pathology	3.35714	1.46322	.208	-.9070	7.6213
		Calcific tendinitis	1.71429	2.01101	.957	-4.1462	7.5748
	rotator cuff tear	impingement	-1.62338	1.27357	.798	-5.3348	2.0881
		G. Arthritis	-6.90909*	2.05883	.014	-12.9090	-.9092
		Instability	-1.29144	1.44353	.947	-5.4982	2.9153
		ACJ pathology	1.73377	1.52829	.866	-2.7200	6.1875
		Calcific tendinitis	.09091	2.05883	1.000	-5.9090	6.0908
	G. Arthritis	impingement	5.28571	2.01101	.101	-.5748	11.1462
		rotator cuff tear	6.90909*	2.05883	.014	.9092	12.9090
		Instability	5.61765	2.12272	.097	-.5684	11.8037
		ACJ pathology	8.64286*	2.18125	.002	2.2863	14.9995
		Calcific tendinitis	7.00000	2.58088	.083	-.5212	14.5212
	Instability	impingement	-.33193	1.37446	1.000	-4.3374	3.6735
		rotator cuff tear	1.29144	1.44353	.947	-2.9153	5.4982
		G. Arthritis	-5.61765	2.12272	.097	-11.8037	.5684
		ACJ pathology	3.02521	1.61332	.424	-1.6764	7.7268

	Calcific tendinitis	1.38235	2.12272	.987	-4.8037	7.5684
ACJ pathology	impingement	-3.35714	1.46322	.208	-7.6213	.9070
	rotator cuff tear	-1.73377	1.52829	.866	-6.1875	2.7200
	G. Arthritis	-8.64286*	2.18125	.002	-14.9995	-2.2863
	Instability	-3.02521	1.61332	.424	-7.7268	1.6764
	Calcific tendinitis	-1.64286	2.18125	.974	-7.9995	4.7137
Calcific tendinitis	impingement	-1.71429	2.01101	.957	-7.5748	4.1462
	rotator cuff tear	-.09091	2.05883	1.000	-6.0908	5.9090
	G. Arthritis	-7.00000	2.58088	.083	-14.5212	.5212
	Instability	-1.38235	2.12272	.987	-7.5684	4.8037
	ACJ pathology	1.64286	2.18125	.974	-4.7137	7.9995

*. The mean difference is significant at the 0.05 level.

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Severity of Pain	.393	5	87	.853
Number of the areas	3.384	5	87	.008

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Severity of Pain	Between Groups	17.856	5	3.571	1.138	.347
	Within Groups	273.068	87	3.139		
	Total	290.925	92			

Number of the areas	Between Groups	354.736	5	70.947	3.550	.006
	Within Groups	1738.511	87	19.983		
	Total	2093.247	92			

Homogeneous Subsets

Severity of Pain

Tukey HSD

shoulder diseases	N	Subset for alpha = 0.05 1
ACJ pathology	14	6.4286
Instability	18	6.7222
rotator cuff tear	22	7.0455
Calcific tendinitis	6	7.5000
impingement	27	7.5185
G. Arthritis	6	7.8333
Sig.		.431

Means for groups in homogeneous subsets are displayed.

Number of the areas

Tukey HSD

shoulder diseases	N	Subset for alpha = 0.05 1	 2
ACJ pathology	14	3.8571	
Calcific tendinitis	6	5.5000	
rotator cuff tear	22	5.5909	
Instability	17	6.8824	
impingement	28	7.2143	7.2143
G. Arthritis	6		12.5000
Sig.		.495	.071

Means for groups in homogeneous subsets are displayed.

F 3.550

, $F(2, 128) = 13.733$, p= .001, $\hat{\eta}_p^2 = .177$

F (5, 87)= 3.550, p=0.006 showed significant difference between the number of the areas marked by the patients. On post hoc tests after finding a significant p value, tukey test was performed and p was > 0.05(0.495) showing the homogeneity of the groups for the number of the areas

Number of the areas

Tukey HSD

shoulder diseases	N	Subset for alpha = 0.05	
		1	2
ACJ pathology	14	3.8571	
Calcific tendinitis	6	5.5000	
rotator cuff tear	22	5.5909	
Instability	17	6.8824	
impingement	28	7.2143	7.2143
G. Arthritis	6		12.5000
Sig.		.495	.071

Means for groups in homogeneous subsets are displayed.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Severity of Pain	Between Groups	17.856	5	3.571	1.138	.347
	Within Groups	273.068	87	3.139		
	Total	290.925	92			
Number of the areas	Between Groups	354.736	5	70.947	3.550	.006
	Within Groups	1738.511	87	19.983		
	Total	2093.247	92			

Post hoc Multiple comparisons showed meaningful difference especially between G arthritis and rotator cuff tears (p= 0.014) and between G arthritis and ACJ pathology (p= 0.002)

Statistics Reports of the Second Phase

Second phase: Agreement of estimation and diagnosis

Crosstabs

Row variable (first classifier): **Overal best-L**
Column variable (second classifier): **Actual**

	1	2	3	4	5	6	7
1	26	3	0	7	0	6	0
2	0	51	2	7	2	3	0
3	1	1	5	1	2	1	0
4	1	5	0	9	2	0	1
5	2	2	0	2	0	5	0
6	0	7	0	0	3	31	0
7	0	1	0	1	2	0	2

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 63.92%
Expected agreement = 22.31%
Kappa = 0.535536 (se = 0.035833)
95% confidence interval = 0.465305 to 0.605766
z (for k = 0) = 14.945507
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:

1	0.8333	0.6667	0.5	0.3333	0.1667	0
0.8333	1	0.8333	0.6667	0.5	0.3333	0.1667
0.6667	0.8333	1	0.8333	0.6667	0.5	0.3333
0.5	0.6667	0.8333	1	0.8333	0.6667	0.5
0.3333	0.5	0.6667	0.8333	1	0.8333	0.6667
0.1667	0.3333	0.5	0.6667	0.8333	1	0.8333
0	0.1667	0.3333	0.5	0.6667	0.8333	1

Observed agreement = 84.45%
Expected agreement = 64.46%
Kappa = 0.562519 (se = 0.052614)
95% confidence interval for kappa = 0.459398 to 0.66564
z (for kw = 0) = 10.691495
P < 0.0001

Scott's pi
Observed agreement = 63.92%
Expected agreement = 22.51%
Pi = 0.534341

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 12.385919 df = 6 P = 0.0539

Symmetry (generalised McNemar) chi-square = 26.266667 df = 21 P = 0.1965

Second Phase: Sensitivity, specificity

Disease- 1

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Tue Oct 8 21:25:52 UTC+0100 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	4	26	30
Test Negative	148	16	164
Totals	152	42	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.216495	0.162062	0.28242
Sensitivity	0.619048	0.456502	0.760097
Specificity	0.973684	0.929772	0.991537
For any particular test result, the probability that it will be:			
Positive	0.154639	0.108328	0.215029
Negative	0.845361	0.784971	0.891672
For any particular positive test result, the probability that it is:			
True Positive	0.866667	0.683577	0.956403
False Positive	0.133333	0.043597	0.316423
For any particular negative test result, the probability that it is:			
True Negative	0.902439	0.843832	0.941427
False Negative	0.097561	0.058573	0.156168
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	23.52381	8.691416	63.66852
Negative [C]	0.391248	0.265994	0.575485
Positive [W]	6.5	2.582498	16.360131
Negative [W]	0.108108	0.067787	0.172414

Disease-2

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Tue Oct 8 21:29:11 UTC+0100 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	19	51	70
Test Negative	110	14	124
Totals	129	65	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.335052	0.270002	0.406753
Sensitivity	0.784615	0.661922	0.8732
Specificity	0.852713	0.777006	0.906764
For any particular test result, the probability that it will be:			
Positive	0.360825	0.29413	0.433128
Negative	0.639175	0.566872	0.70587
For any particular positive test result, the probability that it is:			
True Positive	0.728571	0.607013	0.824771
False Positive	0.271429	0.175229	0.392987
For any particular negative test result, the probability that it is:			
True Negative	0.887097	0.814674	0.934634
False Negative	0.112903	0.065366	0.185326
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	5.327126	3.450437	8.224542
Negative [C]	0.252587	0.158403	0.402774
Positive [W]	2.684211	1.782142	4.04288
Negative [W]	0.127273	0.077553	0.208869

Disease-3

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Tue Oct 8 21:31:24 UTC+0100 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	2	5	7
Test Negative	182	5	187
Totals	184	10	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.051546	0.026406	0.095476
Sensitivity	0.5	0.201423	0.798577
Specificity	0.98913	0.957152	0.998115
For any particular test result, the probability that it will be:			
Positive	0.036082	0.015902	0.075935
Negative	0.963918	0.924065	0.984098
For any particular positive test result, the probability that it is:			
True Positive	0.714286	0.302561	0.948876
False Positive	0.285714	0.051124	0.697439
For any particular negative test result, the probability that it is:			
True Negative	0.973262	0.935341	0.990115
False Negative	0.026738	0.009885	0.064659
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	46	10.148789	208.497789
Negative [C]	0.505495	0.271952	0.939594
Positive [W]	2.5	0.708052	8.827039
Negative [W]	0.027473	0.011567	0.065248

Disease-4

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Tue Oct 8 21:53:14 UTC+0100 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	8	9	17
Test Negative	159	18	177
Totals	167	27	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.139175	0.095283	0.197807
Sensitivity	0.333333	0.172353	0.539849
Specificity	0.952096	0.904533	0.977555
For any particular test result, the probability that it will be:			
Positive	0.087629	0.053411	0.138867
Negative	0.912371	0.861133	0.946589
For any particular positive test result, the probability that it is:			
True Positive	0.529412	0.285339	0.761427
False Positive	0.470588	0.238573	0.714661
For any particular negative test result, the probability that it is:			
True Negative	0.898305	0.841706	0.936939
False Negative	0.101695	0.063061	0.158294
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	6.958333	2.940863	16.464012
Negative [C]	0.70021	0.535967	0.914783
Positive [W]	1.125	0.573035	2.208633
Negative [W]	0.113208	0.072975	0.17562

Disease –6

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Tue Oct 8 21:56:48 UTC+0100 2013)
Values entered:

	Condition		Totals
	Absent	Present	
Test Positive	10	31	41
Test Negative	138	15	153
Totals	148	46	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.237113	0.180425	0.304445
Sensitivity	0.673913	0.518654	0.800308
Specificity	0.932432	0.875946	0.965298
For any particular test result, the probability that it will be:			
Positive	0.21134	0.157503	0.276882
Negative	0.78866	0.723118	0.842497
For any particular positive test result, the probability that it is:			
True Positive	0.756098	0.593558	0.870922
False Positive	0.243902	0.129078	0.406442
For any particular negative test result, the probability that it is:			
True Negative	0.901961	0.840707	0.942205
False Negative	0.098039	0.057795	0.159293
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	9.973913	5.304884	18.75233
Negative [C]	0.349716	0.23061	0.53034
Positive [W]	3.1	1.759661	5.461281
Negative [W]	0.108696	0.067124	0.176013

Disease -7

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Tue Oct 8 21:58:30 UTC+0100 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	1	2	3
Test Negative	187	4	191
Totals	188	6	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.030928	0.012634	0.069224
Sensitivity	0.333333	0.05999	0.758921
Specificity	0.994681	0.966183	0.999722
For any particular test result, the probability that it will be:			
Positive	0.015464	0.004002	0.048173
Negative	0.984536	0.951827	0.995998
For any particular positive test result, the probability that it is:			
True Positive	0.666667	0.125335	0.982347
False Positive	0.333333	0.017653	0.874665
For any particular negative test result, the probability that it is:			
True Negative	0.979058	0.94377	0.99327
False Negative	0.020942	0.00673	0.05623
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	62.666667	6.548144	599.728909
Negative [C]	0.670232	0.38061	1.180238
Positive [W]	2	0.334191	11.969212
Negative [W]	0.02139	0.00811	0.05642

Second Phase: Disease group- Map Group Agreement

Crosstabs

Row variable (first classifier): **Disease group**
Column variable (second classifier): **map group**

	1	2
1	71	22
2	13	88

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 81.96%
Expected agreement = 50.28%
Kappa = 0.63717 (se = 0.071483)
95% confidence interval = 0.497067 to 0.777273
z (for k = 0) = 8.913632
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:
1 0
0 1
Observed agreement = 81.96%
Expected agreement = 50.28%
Kappa = 0.63717 (se = 0.071483)
95% confidence interval for kappa = 0.497067 to 0.777273
z (for kw = 0) = 8.913632
P < 0.0001

Scott's pi
Observed agreement = 81.96%
Expected agreement = 50.38%
Pi = 0.636383
95% Donner-Eliasziw confidence interval = 0.515375 to 0.732535

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 2.314286 df = 1 P = 0.1282

Symmetry (generalised McNemar) chi-square = 2.314286 df = 1 P = 0.1282

Second Phase: Group Estimation – Sensitivity, Specificity

For Group-1 Shoulder Disease:

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Tue Oct 8 21:18:00 UTC+0100 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	13	71	84
Test Negative	88	22	110
Totals	101	93	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.479381	0.407656	0.551937
Sensitivity	0.763441	0.661921	0.842795
Specificity	0.871287	0.786408	0.926978
For any particular test result, the probability that it will be:			
Positive	0.43299	0.362756	0.505921
Negative	0.56701	0.494079	0.637244
For any particular positive test result, the probability that it is:			
True Positive	0.845238	0.746225	0.911827
False Positive	0.154762	0.088173	0.253775
For any particular negative test result, the probability that it is:			
True Negative	0.8	0.71072	0.867834
False Negative	0.2	0.132166	0.28928
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	5.931348	3.526781	9.975354
Negative [C]	0.271505	0.187918	0.392273
Positive [W]	5.461538	3.285958	9.077537
Negative [W]	0.25	0.171273	0.364915

For Group -2 Diseases

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Tue Oct 8 21:20:01 UTC+0100 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	22	88	110
Test Negative	71	13	84
Totals	93	101	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.520619	0.448063	0.592344
Sensitivity	0.871287	0.786408	0.926978
Specificity	0.763441	0.661921	0.842795
For any particular test result, the probability that it will be:			
Positive	0.56701	0.494079	0.637244
Negative	0.43299	0.362756	0.505921
For any particular positive test result, the probability that it is:			
True Positive	0.8	0.71072	0.867834
False Positive	0.2	0.132166	0.28928
For any particular negative test result, the probability that it is:			
True Negative	0.845238	0.746225	0.911827
False Negative	0.154762	0.088173	0.253775
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	3.683168	2.537168	5.3468
Negative [C]	0.168596	0.100843	0.281868
Positive [W]	4	2.721116	5.879941
Negative [W]	0.183099	0.110608	0.303099

Third Phase

Estimation of the Disease Groups by Testers

Crosstabs

Row variable (first classifier): **A-1**
Column variable (second classifier): **Disease group**

	1	2
1	74	29
2	19	72

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 75.26%
Expected agreement = 49.87%
Kappa = 0.506414 (se = 0.071415)
95% confidence interval = 0.366442 to 0.646385
z (for k = 0) = 7.091117
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:
1 0
0 1
Observed agreement = 75.26%
Expected agreement = 49.87%
Kappa = 0.506414 (se = 0.071415)
95% confidence interval for kappa = 0.366442 to 0.646385
z (for kw = 0) = 7.091117
P < 0.0001

Scott's pi
Observed agreement = 75.26%
Expected agreement = 50.01%
Pi = 0.505102
95% Donner-Eliasziw confidence interval = 0.374627 to 0.615961

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 2.083333 df = 1 P = 0.1489

Symmetry (generalised McNemar) chi-square = 2.083333 df = 1 P = 0.1489

Crosstabs

Row variable (first classifier): **B-1**
Column variable (second classifier): **Disease group**

	1	2
1	77	20
2	16	81

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 81.44%
Expected agreement = 50%
Kappa = 0.628866 (se = 0.071735)
95% confidence interval = 0.488268 to 0.769463
z (for k = 0) = 8.766546
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:
1 0
0 1
Observed agreement = 81.44%
Expected agreement = 50%
Kappa = 0.628866 (se = 0.071735)
95% confidence interval for kappa = 0.488268 to 0.769463
z (for kw = 0) = 8.766546
P < 0.0001

Scott's pi
Observed agreement = 81.44%
Expected agreement = 50.02%
Pi = 0.628708
95% Donner-Eliasziw confidence interval = 0.50743 to 0.725563

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 0.444444 df = 1 P = 0.505

Symmetry (generalised McNemar) chi-square = 0.444444 df = 1 P = 0.505

Crosstabs

Row variable (first classifier): **C-1**
Column variable (second classifier): **Disease group**

	1	**2**
1	71	15
2	22	86

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 80.93%
Expected agreement = 50.23%
Kappa = 0.616765 (se = 0.071607)
95% confidence interval = 0.476418 to 0.757111
z (for k = 0) = 8.613205
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:
1 0
0 1

Observed agreement = 80.93%
Expected agreement = 50.23%
Kappa = 0.616765 (se = 0.071607)
95% confidence interval for kappa = 0.476418 to 0.757111
z (for kw = 0) = 8.613205
P < 0.0001

Scott's pi
Observed agreement = 80.93%
Expected agreement = 50.3%
Pi = 0.616263
95% Donner-Eliasziw confidence interval = 0.493553 to 0.714938

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 1.324324 df = 1 P = 0.2498

Symmetry (generalised McNemar) chi-square = 1.324324 df = 1 P = 0.2498

Third Phase: Disease Groups: Sensitivity, Specificity

Rater A

Group -1

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 12:36:30 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	29	74	103
Test Negative	72	19	91
Totals	101	93	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.479381	0.407656	0.551937
Sensitivity	0.795699	0.696916	0.86953
Specificity	0.712871	0.612929	0.796359
For any particular test result, the probability that it will be:			
Positive	0.530928	0.458238	0.602373
Negative	0.469072	0.397627	0.541762
For any particular positive test result, the probability that it is:			
True Positive	0.718447	0.619847	0.80048
False Positive	0.281553	0.19952	0.380153
For any particular negative test result, the probability that it is:			
True Negative	0.791209	0.690754	0.866568
False Negative	0.208791	0.133432	0.309246
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	2.771227	2.004108	3.831979
Negative [C]	0.286589	0.190403	0.431366
Positive [W]	2.551724	1.832045	3.554114
Negative [W]	0.263889	0.175895	0.395904

Rater A

Group -2

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 12:37:31 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	19	72	91
Test Negative	74	29	103
Totals	93	101	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.520619	0.448063	0.592344
Sensitivity	0.712871	0.612929	0.796359
Specificity	0.795699	0.696916	0.86953
For any particular test result, the probability that it will be:			
Positive	0.469072	0.397627	0.541762
Negative	0.530928	0.458238	0.602373
For any particular positive test result, the probability that it is:			
True Positive	0.791209	0.690754	0.866568
False Positive	0.208791	0.133432	0.309246
For any particular negative test result, the probability that it is:			
True Negative	0.718447	0.619847	0.80048
False Negative	0.281553	0.19952	0.380153
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	3.489317	2.293206	5.309309
Negative [C]	0.360851	0.26395	0.493326
Positive [W]	3.789474	2.50572	5.730933
Negative [W]	0.391892	0.285727	0.537504

Rater B

Group -1

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 12:31:48 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	20	77	97
Test Negative	81	16	97
Totals	101	93	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.479381	0.407656	0.551937
Sensitivity	0.827957	0.73263	0.895504
Specificity	0.80198	0.708391	0.872071
For any particular test result, the probability that it will be:			
Positive	0.5	0.427802	0.572198
Negative	0.5	0.427802	0.572198
For any particular positive test result, the probability that it is:			
True Positive	0.793814	0.697226	0.866626
False Positive	0.206186	0.133374	0.302774
For any particular negative test result, the probability that it is:			
True Negative	0.835052	0.742928	0.899927
False Negative	0.164948	0.100073	0.257072
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	4.181183	2.7936	6.257979
Negative [C]	0.214523	0.136698	0.336655
Positive [W]	3.85	2.571895	5.763261
Negative [W]	0.197531	0.125731	0.310333

Rater B

Group -2

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 12:34:40 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	16	81	97
Test Negative	77	20	97
Totals	93	101	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.520619	0.448063	0.592344
Sensitivity	0.80198	0.708391	0.872071
Specificity	0.827957	0.73263	0.895504
For any particular test result, the probability that it will be:			
Positive	0.5	0.427802	0.572198
Negative	0.5	0.427802	0.572198
For any particular positive test result, the probability that it is:			
True Positive	0.835052	0.742928	0.899927
False Positive	0.164948	0.100073	0.257072
For any particular negative test result, the probability that it is:			
True Negative	0.793814	0.697226	0.866626
False Negative	0.206186	0.133374	0.302774
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	4.66151	2.953754	7.35663
Negative [C]	0.239167	0.160823	0.355675
Positive [W]	5.0625	3.207366	7.990639
Negative [W]	0.25974	0.174856	0.385833

Rater C

Group -1

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 12:39:11 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	15	71	86
Test Negative	86	22	108
Totals	101	93	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.479381	0.407656	0.551937
Sensitivity	0.763441	0.661921	0.842795
Specificity	0.851485	0.763685	0.911754
For any particular test result, the probability that it will be:			
Positive	0.443299	0.372682	0.516198
Negative	0.556701	0.483802	0.627318
For any particular positive test result, the probability that it is:			
True Positive	0.825581	0.725445	0.895931
False Positive	0.174419	0.104069	0.274555
For any particular negative test result, the probability that it is:			
True Negative	0.796296	0.705719	0.865311
False Negative	0.203704	0.134689	0.294281
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	5.140502	3.17933	8.311423
Negative [C]	0.277819	0.192189	0.401604
Positive [W]	4.733333	2.958454	7.573025
Negative [W]	0.255814	0.175375	0.373148

Rater C

Group -2

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 12:40:02 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	22	86	108
Test Negative	71	15	86
Totals	93	101	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.520619	0.448063	0.592344
Sensitivity	0.851485	0.763685	0.911754
Specificity	0.763441	0.661921	0.842795
For any particular test result, the probability that it will be:			
Positive	0.556701	0.483802	0.627318
Negative	0.443299	0.372682	0.516198
For any particular positive test result, the probability that it is:			
True Positive	0.796296	0.705719	0.865311
False Positive	0.203704	0.134689	0.294281
For any particular negative test result, the probability that it is:			
True Negative	0.825581	0.725445	0.895931
False Negative	0.174419	0.104069	0.274555
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	3.59946	2.476137	5.23239
Negative [C]	0.194534	0.121159	0.312344
Positive [W]	3.909091	2.660233	5.74423
Negative [W]	0.211268	0.132754	0.336216

Third Phase: Group Agreement between the Testers

General agreement over all (2) categories with 3 raters per subject

Cohen's kappa (Fleiss-Cuzick extension):

Kappa = 0.697505 (se = 0.041451)
z (for k = 0) = 16.827087
P < 0.0001

General agreement over all (2) categories with 3 raters per subject

Cohen's kappa (Fleiss-Cuzick extension):

Kappa = 0.697505 (se = 0.041451)
z (for k = 0) = 16.827087
P < 0.0001

Crosstabs

Row variable (first classifier): **A-1**
Column variable (second classifier): **B-1**

	1	2
1	83	20
2	14	77

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 82.47%
Expected agreement = 50%
Kappa = 0.649485 (se = 0.071658)
95% confidence interval = 0.509037 to 0.789932
z (for k = 0) = 9.063629
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:

1	0
0	1

Observed agreement = 82.47%
Expected agreement = 50%
Kappa = 0.649485 (se = 0.071658)
95% confidence interval for kappa = 0.509037 to 0.789932
z (for kw = 0) = 9.063629
P < 0.0001

Scott's pi
Observed agreement = 82.47%
Expected agreement = 50.05%
Pi = 0.649149
95% Donner-Eliasziw confidence interval = 0.529747 to 0.743323

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 1.058824 df = 1 P = 0.3035

Symmetry (generalised McNemar) chi-square = 1.058824 df = 1 P = 0.3035

Crosstabs

Row variable (first classifier): **A-1**
Column variable (second classifier): **C-1**

	1	2
1	77	26
2	9	82

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 81.96%
Expected agreement = 49.65%
Kappa = 0.641689 (se = 0.0707)
95% confidence interval = 0.503119 to 0.780258
z (for k = 0) = 9.076199
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:

1	0
0	1

Observed agreement = 81.96%
Expected agreement = 49.65%
Kappa = 0.641689 (se = 0.0707)
95% confidence interval for kappa = 0.503119 to 0.780258
z (for kw = 0) = 9.076199
P < 0.0001

Scott's pi
Observed agreement = 81.96%
Expected agreement = 50.03%
Pi = 0.638935
95% Donner-Eliasziw confidence interval = 0.518582 to 0.734464

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 8.257143 df = 1 P = 0.0041

Symmetry (generalised McNemar) chi-square = 8.257143 df = 1 P = 0.0041

Crosstabs

Row variable (first classifier): **B-1**
Column variable (second classifier): **C-1**

	1	2
1	82	15
2	4	93

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 90.21%
Expected agreement = 50%
Kappa = 0.804124 (se = 0.071333)
95% confidence interval = 0.664314 to 0.943933
z (for k = 0) = 11.272867
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:
1 0
0 1
Observed agreement = 90.21%
Expected agreement = 50%
Kappa = 0.804124 (se = 0.071333)
95% confidence interval for kappa = 0.664314 to 0.943933
z (for kw = 0) = 11.272867
P < 0.0001

Scott's pi
Observed agreement = 90.21%
Expected agreement = 50.16%
Pi = 0.803492
95% Donner-Eliasziw confidence interval = 0.703301 to 0.872387

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 6.368421 df = 1 P = 0.0116

Symmetry (generalised McNemar) chi-square = 6.368421 df = 1 P = 0.0116

Third Phase: Estimation of Individual Shoulder Disorders by Testers

Crosstabs

Row variable (first classifier): **Overall best Est-A**
Column variable (second classifier): **Actual**

	1	2	3	4	5	6	7
1	16	3	0	6	1	7	0
2	7	42	0	1	2	5	0
3	0	1	7	2	1	1	0
4	4	8	0	16	1	5	2
5	1	0	0	2	12	5	0
6	1	3	0	1	5	23	0
7	0	0	0	1	0	0	2

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 60.82%
Expected agreement = 19.41%
Kappa = 0.513913 (se = 0.033852)
95% confidence interval = 0.447564 to 0.580261
z (for k = 0) = 15.181235
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:

1	0.8333	0.6667	0.5	0.3333	0.1667	0
0.8333	1	0.8333	0.6667	0.5	0.3333	0.1667
0.6667	0.8333	1	0.8333	0.6667	0.5	0.3333
0.5	0.6667	0.8333	1	0.8333	0.6667	0.5
0.3333	0.5	0.6667	0.8333	1	0.8333	0.6667
0.1667	0.3333	0.5	0.6667	0.8333	1	0.8333
0	0.1667	0.3333	0.5	0.6667	0.8333	1

Observed agreement = 84.02%
Expected agreement = 65.09%
Kappa = 0.542209 (se = 0.050451)
95% confidence interval for kappa = 0.443326 to 0.641092
z (for kw = 0) = 10.747145
P < 0.0001

Scott's pi
Observed agreement = 60.82%
Expected agreement = 19.58%
Pi = 0.512853

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 11.975026 df = 6 P = 0.0625

Symmetry (generalised McNemar) chi-square = 22.777778 df = 21 P = 0.3559

Crosstabs

Row variable (first classifier): **Overall best Est-B**

Column variable (second classifier): **Actual**

	1	2	3	4	5	6	7
1	21	8	0	5	3	4	0
2	4	46	0	0	0	0	0
3	0	1	3	0	0	0	0
4	2	1	2	21	3	9	1
5	1	0	1	2	7	6	0
6	1	1	1	1	6	25	2
7	0	0	0	0	3	2	1

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 63.92%
Expected agreement = 19.39%
Kappa = 0.552377 (se = 0.034078)
95% confidence interval = 0.485585 to 0.619168
z (for k = 0) = 16.209175
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:

1	0.8333	0.6667	0.5	0.3333	0.1667	0
0.8333	1	0.8333	0.6667	0.5	0.3333	0.1667
0.6667	0.8333	1	0.8333	0.6667	0.5	0.3333
0.5	0.6667	0.8333	1	0.8333	0.6667	0.5
0.3333	0.5	0.6667	0.8333	1	0.8333	0.6667
0.1667	0.3333	0.5	0.6667	0.8333	1	0.8333
0	0.1667	0.3333	0.5	0.6667	0.8333	1

Observed agreement = 88.14%
Expected agreement = 64%
Kappa = 0.670669 (se = 0.051021)
95% confidence interval for kappa = 0.570669 to 0.770669
z (for kw = 0) = 13.14484
P < 0.0001

Scott's pi
Observed agreement = 63.92%
Expected agreement = 19.66%
Pi = 0.550852

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 13.67632 df = 6 P = 0.0335

Symmetry (generalised McNemar) chi-square = 23.019048 df = 21 P = 0.343

Crosstabs

Row variable (first classifier): **overall best Est-C**
Column variable (second classifier): **Actual**

	1	2	3	4	5	6	7
1	16	3	0	3	1	5	0
2	5	37	0	1	0	0	0
3	2	3	5	3	1	1	0
4	2	10	1	15	0	13	0

5	3	1	1	0	12	3	0
6	0	3	0	4	1	24	0
7	1	0	0	3	7	0	4

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 58.25%
Expected agreement = 17.35%
Kappa = 0.49484 (se = 0.031743)
95% confidence interval = 0.432626 to 0.557055
z (for k = 0) = 15.58912
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:

1	0.8333	0.6667	0.5	0.3333	0.1667	0
0.8333	1	0.8333	0.6667	0.5	0.3333	0.1667
0.6667	0.8333	1	0.8333	0.6667	0.5	0.3333
0.5	0.6667	0.8333	1	0.8333	0.6667	0.5
0.3333	0.5	0.6667	0.8333	1	0.8333	0.6667
0.1667	0.3333	0.5	0.6667	0.8333	1	0.8333
0	0.1667	0.3333	0.5	0.6667	0.8333	1

Observed agreement = 84.02%
Expected agreement = 63.95%
Kappa = 0.556773 (se = 0.049343)
95% confidence interval for kappa = 0.460063 to 0.653483
z (for kw = 0) = 11.283786
P < 0.0001

Scott's pi
Observed agreement = 58.25%
Expected agreement = 17.83%
Pi = 0.491876

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 31.465573 df = 6 P < 0.0001

Symmetry (generalised McNemar) chi-square = 41.828342 df = 21 P = 0.0044

Third Phase: Sensitivity, Specificity of Estimations of Individual

Disorders by the Testers

Rater A Disease 1
VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:34:14 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	17	16	33
Test Negative	148	13	161
Totals	165	29	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.149485	0.103959	0.209308
Sensitivity	0.551724	0.359805	0.73046
Specificity	0.89697	0.837675	0.937056
For any particular test result, the probability that it will be:			
Positive	0.170103	0.121545	0.232086
Negative	0.829897	0.767914	0.878455
For any particular positive test result, the probability that it is:			
True Positive	0.484848	0.311668	0.661454
False Positive	0.515152	0.338546	0.688332
For any particular negative test result, the probability that it is:			
True Negative	0.919255	0.863043	0.954538
False Negative	0.080745	0.045462	0.136957
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	5.35497	3.067823	9.347248
Negative [C]	0.499767	0.333287	0.749404
Positive [W]	0.941176	0.580668	1.525508
Negative [W]	0.087838	0.052104	0.148079

Rater A

Disease 2

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:36:26 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	15	42	57
Test Negative	122	15	137
Totals	137	57	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.293814	0.231852	0.364102
Sensitivity	0.736842	0.600941	0.840633
Specificity	0.890511	0.822942	0.935341
For any particular test result, the probability that it will be:			
Positive	0.293814	0.231852	0.364102
Negative	0.706186	0.635898	0.768148
For any particular positive test result, the probability that it is:			
True Positive	0.736842	0.600941	0.840633
False Positive	0.263158	0.159367	0.399059
For any particular negative test result, the probability that it is:			
True Negative	0.890511	0.822942	0.935341
False Negative	0.109489	0.064659	0.177058
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	6.729825	4.073196	11.119166
Negative [C]	0.295513	0.191054	0.457086
Positive [W]	2.8	1.765348	4.441052
Negative [W]	0.122951	0.076132	0.198561

Rater A

Disease 3

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:38:18 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	5	7	12
Test Negative	182	0	182
Totals	187	7	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.036082	0.015902	0.075935
Sensitivity	1	0.560935	1
Specificity	0.973262	0.935341	0.990115
For any particular test result, the probability that it will be:			
Positive	0.061856	0.033836	0.108132
Negative	0.938144	0.891868	0.966164
For any particular positive test result, the probability that it is:			
True Positive	0.583333	0.285993	0.835007
False Positive	0.416667	0.164993	0.714007
For any particular negative test result, the probability that it is:			
True Negative	1	0.974235	1
False Negative	0	0	0.025765
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	37.4	15.751688	88.800642
Negative [C]	0	0	NaN
Positive [W]	1.4	0.614945	3.187275
Negative [W]	0	0	NaN

Rater A

Disease 4

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:29:16 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	22	16	38
Test Negative	143	13	156
Totals	165	29	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.149485	0.103959	0.209308
Sensitivity	0.551724	0.359805	0.73046
Specificity	0.866667	0.802943	0.912755
For any particular test result, the probability that it will be:			
Positive	0.195876	0.143908	0.260192
Negative	0.804124	0.739808	0.856092
For any particular positive test result, the probability that it is:			
True Positive	0.421053	0.267203	0.590555
False Positive	0.578947	0.409445	0.732797
For any particular negative test result, the probability that it is:			
True Negative	0.916667	0.85882	0.953062
False Negative	0.083333	0.046938	0.14118
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	4.137931	2.4876	6.883129
Negative [C]	0.517241	0.344781	0.775967
Positive [W]	0.727273	0.458657	1.153205
Negative [W]	0.090909	0.053961	0.153156

Rater A

Disease 5

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:40:25 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	8	12	20
Test Negative	164	10	174
Totals	172	22	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.113402	0.073991	0.168673
Sensitivity	0.545455	0.32674	0.749293
Specificity	0.953488	0.907225	0.978213
For any particular test result, the probability that it will be:			
Positive	0.103093	0.065661	0.156842
Negative	0.896907	0.843158	0.934339
For any particular positive test result, the probability that it is:			
True Positive	0.6	0.364117	0.800229
False Positive	0.4	0.199771	0.635883
For any particular negative test result, the probability that it is:			
True Negative	0.942529	0.893896	0.970531
False Negative	0.057471	0.029469	0.106104
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	11.727273	5.3933	25.499959
Negative [C]	0.476718	0.301455	0.753877
Positive [W]	1.5	0.786914	2.859272
Negative [W]	0.060976	0.033384	0.111373

Rater A

Disease 6

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:42:02 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	10	23	33
Test Negative	128	23	151
Totals	138	46	184

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.25	0.190538	0.320126
Sensitivity	0.5	0.351188	0.648812
Specificity	0.927536	0.867314	0.962754
For any particular test result, the probability that it will be:			
Positive	0.179348	0.128308	0.244134
Negative	0.820652	0.755866	0.871692
For any particular positive test result, the probability that it is:			
True Positive	0.69697	0.511269	0.837871
False Positive	0.30303	0.162129	0.488731
For any particular negative test result, the probability that it is:			
True Negative	0.847682	0.778094	0.899045
False Negative	0.152318	0.100955	0.221906
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	6.9	3.554925	13.392686
Negative [C]	0.539063	0.403325	0.720483
Positive [W]	2.3	1.308243	4.043591
Negative [W]	0.179688	0.123053	0.262387

Rater A

Disease 7

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:43:09 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	1	2	3
Test Negative	189	2	191
Totals	190	4	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.020619	0.006625	0.055381
Sensitivity	0.5	0.091899	0.908101
Specificity	0.994737	0.966531	0.999725
For any particular test result, the probability that it will be:			
Positive	0.015464	0.004002	0.048173
Negative	0.984536	0.951827	0.995998
For any particular positive test result, the probability that it is:			
True Positive	0.666667	0.125335	0.982347
False Positive	0.333333	0.017653	0.874665
For any particular negative test result, the probability that it is:			
True Negative	0.989529	0.958688	0.998184
False Negative	0.010471	0.001816	0.041312
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	95	10.667408	846.034936
Negative [C]	0.502646	0.188642	1.339323
Positive [W]	2	0.334191	11.969212
Negative [W]	0.010582	0.002666	0.042009

Rater B

Disease 1

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:46:20 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	6	21	27
Test Negative	147	20	167
Totals	153	41	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.21134	0.157503	0.276882
Sensitivity	0.512195	0.353655	0.668486
Specificity	0.960784	0.912816	0.983959
For any particular test result, the probability that it will be:			
Positive	0.139175	0.095283	0.197807
Negative	0.860825	0.802193	0.904717
For any particular positive test result, the probability that it is:			
True Positive	0.777778	0.572664	0.906244
False Positive	0.222222	0.093756	0.427336
For any particular negative test result, the probability that it is:			
True Negative	0.88024	0.81883	0.923537
False Negative	0.11976	0.076463	0.18117
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	13.060976	5.642708	30.231774
Negative [C]	0.507715	0.37083	0.695129
Positive [W]	3.5	1.68011	7.291191
Negative [W]	0.136054	0.090042	0.205579

Rater B

Disease 2

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:47:37 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	4	46	50
Test Negative	133	11	144
Totals	137	57	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.293814	0.231852	0.364102
Sensitivity	0.807018	0.676833	0.895278
Specificity	0.970803	0.922337	0.990606
For any particular test result, the probability that it will be:			
Positive	0.257732	0.198976	0.326285
Negative	0.742268	0.673715	0.801024
For any particular positive test result, the probability that it is:			
True Positive	0.92	0.7989	0.97406
False Positive	0.08	0.02594	0.2011
For any particular negative test result, the probability that it is:			
True Negative	0.923611	0.864183	0.959341
False Negative	0.076389	0.040659	0.135817
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	27.640351	10.437441	73.196966
Negative [C]	0.198786	0.116854	0.338166
Positive [W]	11.5	4.476476	29.543326
Negative [W]	0.082707	0.046817	0.14611

Rater B

Disease 3

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:48:34 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	1	3	4
Test Negative	186	4	190
Totals	187	7	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.036082	0.015902	0.075935
Sensitivity	0.428571	0.118083	0.797628
Specificity	0.994652	0.966006	0.999721
For any particular test result, the probability that it will be:			
Positive	0.020619	0.006625	0.055381
Negative	0.979381	0.944619	0.993375
For any particular positive test result, the probability that it is:			
True Positive	0.75	0.219427	0.986809
False Positive	0.25	0.013191	0.780573
For any particular negative test result, the probability that it is:			
True Negative	0.978947	0.943481	0.993235
False Negative	0.021053	0.006765	0.056519
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	80.142857	9.488946	676.879998
Negative [C]	0.574501	0.302444	1.091279
Positive [W]	3	0.501286	17.953818
Negative [W]	0.021505	0.008154	0.05672

Rater B

Disease 4

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:50:27 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	19	21	40
Test Negative	146	8	154
Totals	165	29	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.149485	0.103959	0.209308
Sensitivity	0.724138	0.525149	0.865541
Specificity	0.884848	0.823669	0.927457
For any particular test result, the probability that it will be:			
Positive	0.206186	0.152958	0.271332
Negative	0.793814	0.728668	0.847042
For any particular positive test result, the probability that it is:			
True Positive	0.525	0.363442	0.681838
False Positive	0.475	0.318162	0.636558
For any particular negative test result, the probability that it is:			
True Negative	0.948052	0.896746	0.975645
False Negative	0.051948	0.024355	0.103254
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	6.288566	3.895475	10.151797
Negative [C]	0.311762	0.172607	0.563102
Positive [W]	1.105263	0.712287	1.715048
Negative [W]	0.054795	0.027888	0.107662

Rater B

Disease 5

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:51:55 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	10	7	17
Test Negative	162	15	177
Totals	172	22	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.113402	0.073991	0.168673
Sensitivity	0.318182	0.147343	0.548842
Specificity	0.94186	0.892701	0.970185
For any particular test result, the probability that it will be:			
Positive	0.087629	0.053411	0.138867
Negative	0.912371	0.861133	0.946589
For any particular positive test result, the probability that it is:			
True Positive	0.411765	0.194279	0.665465
False Positive	0.588235	0.334535	0.805721
For any particular negative test result, the probability that it is:			
True Negative	0.915254	0.861549	0.950143
False Negative	0.084746	0.049857	0.138451
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	5.472727	2.32074	12.905687
Negative [C]	0.723906	0.543757	0.963739
Positive [W]	0.7	0.349865	1.40054
Negative [W]	0.092593	0.056999	0.150412

Rater B

Disease 6

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:53:37 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	12	25	37
Test Negative	136	21	157
Totals	148	46	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.237113	0.180425	0.304445
Sensitivity	0.543478	0.391558	0.688211
Specificity	0.918919	0.859566	0.955513
For any particular test result, the probability that it will be:			
Positive	0.190722	0.139404	0.2546
Negative	0.809278	0.7454	0.860596
For any particular positive test result, the probability that it is:			
True Positive	0.675676	0.501055	0.814449
False Positive	0.324324	0.185551	0.498945
For any particular negative test result, the probability that it is:			
True Negative	0.866242	0.800554	0.913416
False Negative	0.133758	0.086584	0.199446
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	6.702899	3.665471	12.257319
Negative [C]	0.496803	0.362028	0.681752
Positive [W]	2.083333	1.243698	3.489817
Negative [W]	0.154412	0.103506	0.230354

Rater B

Disease 7

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:54:40 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	5	1	6
Test Negative	185	3	188
Totals	190	4	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.020619	0.006625	0.055381
Sensitivity	0.25	0.013191	0.780573
Specificity	0.973684	0.936336	0.990272
For any particular test result, the probability that it will be:			
Positive	0.030928	0.012634	0.069224
Negative	0.969072	0.930776	0.987366
For any particular positive test result, the probability that it is:			
True Positive	0.166667	0.008762	0.635177
False Positive	0.833333	0.364823	0.991238
For any particular negative test result, the probability that it is:			
True Negative	0.984043	0.950325	0.99587
False Negative	0.015957	0.00413	0.049675
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	9.5	1.413768	63.836516
Negative [C]	0.77027	0.437323	1.356702
Positive [W]	0.2	0.032256	1.240092
Negative [W]	0.016216	0.005277	0.049834

Rater C

Disease 1

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:57:25 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	12	16	28
Test Negative	153	13	166
Totals	165	29	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.149485	0.103959	0.209308
Sensitivity	0.551724	0.359805	0.73046
Specificity	0.927273	0.873527	0.960149
For any particular test result, the probability that it will be:			
Positive	0.14433	0.099611	0.203567
Negative	0.85567	0.796433	0.900389
For any particular positive test result, the probability that it is:			
True Positive	0.571429	0.374319	0.74973
False Positive	0.428571	0.25027	0.625681
For any particular negative test result, the probability that it is:			
True Negative	0.921687	0.867021	0.955924
False Negative	0.078313	0.044076	0.132979
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	7.586207	4.016279	14.329319
Negative [C]	0.483435	0.322535	0.7246
Positive [W]	1.333333	0.781187	2.275739
Negative [W]	0.084967	0.05037	0.143329

Rater C

Disease 2

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 10:58:59 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	6	37	43
Test Negative	131	20	151
Totals	137	57	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.293814	0.231852	0.364102
Sensitivity	0.649123	0.510578	0.76755
Specificity	0.956204	0.902995	0.982073
For any particular test result, the probability that it will be:			
Positive	0.221649	0.166634	0.287944
Negative	0.778351	0.712056	0.833366
For any particular positive test result, the probability that it is:			
True Positive	0.860465	0.713734	0.941972
False Positive	0.139535	0.058028	0.286266
For any particular negative test result, the probability that it is:			
True Negative	0.86755	0.800498	0.915277
False Negative	0.13245	0.084723	0.199502
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	14.821637	6.624087	33.163958
Negative [C]	0.366948	0.257615	0.522682
Positive [W]	6.166667	2.907324	13.079991
Negative [W]	0.152672	0.101302	0.230092

Rater C

Disease 3

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 11:00:15 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	10	5	15
Test Negative	177	2	179
Totals	187	7	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.036082	0.015902	0.075935
Sensitivity	0.714286	0.302561	0.948876
Specificity	0.946524	0.901055	0.972597
For any particular test result, the probability that it will be:			
Positive	0.07732	0.045437	0.126705
Negative	0.92268	0.873295	0.954563
For any particular positive test result, the probability that it is:			
True Positive	0.333333	0.129878	0.613134
False Positive	0.666667	0.386866	0.870122
For any particular negative test result, the probability that it is:			
True Negative	0.988827	0.955983	0.998062
False Negative	0.011173	0.001938	0.044017
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	13.357143	6.224126	28.664791
Negative [C]	0.301856	0.093509	0.974421
Positive [W]	0.5	0.224631	1.112938
Negative [W]	0.011299	0.002848	0.044836

Rater C

Disease 4

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 11:05:46 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	26	15	41
Test Negative	139	14	153
Totals	165	29	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.149485	0.103959	0.209308
Sensitivity	0.517241	0.328996	0.701094
Specificity	0.842424	0.775754	0.892682
For any particular test result, the probability that it will be:			
Positive	0.21134	0.157503	0.276882
Negative	0.78866	0.723118	0.842497
For any particular positive test result, the probability that it is:			
True Positive	0.365854	0.225698	0.530805
False Positive	0.634146	0.469195	0.774302
For any particular negative test result, the probability that it is:			
True Negative	0.908497	0.848403	0.9472
False Negative	0.091503	0.0528	0.151597
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	3.282493	1.994728	5.401619
Negative [C]	0.573059	0.392304	0.837097
Positive [W]	0.576923	0.362295	0.918698
Negative [W]	0.100719	0.061055	0.166153

Rater C

Disease 5

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 11:07:21 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	8	12	20
Test Negative	164	10	174
Totals	172	22	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.113402	0.073991	0.168673
Sensitivity	0.545455	0.32674	0.749293
Specificity	0.953488	0.907225	0.978213
For any particular test result, the probability that it will be:			
Positive	0.103093	0.065661	0.156842
Negative	0.896907	0.843158	0.934339
For any particular positive test result, the probability that it is:			
True Positive	0.6	0.364117	0.800229
False Positive	0.4	0.199771	0.635883
For any particular negative test result, the probability that it is:			
True Negative	0.942529	0.893896	0.970531
False Negative	0.057471	0.029469	0.106104
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	11.727273	5.3933	25.499959
Negative [C]	0.476718	0.301455	0.753877
Positive [W]	1.5	0.786914	2.859272
Negative [W]	0.060976	0.033384	0.111373

Rater C

Disease 6

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 11:08:40 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	8	24	32
Test Negative	140	22	162
Totals	148	46	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.237113	0.180425	0.304445
Sensitivity	0.521739	0.371252	0.668633
Specificity	0.945946	0.892707	0.974649
For any particular test result, the probability that it will be:			
Positive	0.164948	0.117121	0.226418
Negative	0.835052	0.773582	0.882879
For any particular positive test result, the probability that it is:			
True Positive	0.75	0.562496	0.878729
False Positive	0.25	0.121271	0.437504
For any particular negative test result, the probability that it is:			
True Negative	0.864198	0.799452	0.911107
False Negative	0.135802	0.088893	0.200548
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	9.652174	4.658251	19.999882
Negative [C]	0.50559	0.373583	0.684243
Positive [W]	3	1.593667	5.647351
Negative [W]	0.157143	0.106357	0.23218

Rater C

Disease 7

VassarStats Printable Report:
From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values, and likelihood Ratios
Wed Nov 20 11:09:40 UTC 2013)
Values entered:

	Condition		
	Absent	Present	Totals
Test Positive	11	4	15
Test Negative	179	0	179
Totals	190	4	194

	Estimated Value	95% Confidence Interval	
		Lower Limit	Upper Limit
Prevalence	0.020619	0.006625	0.055381
Sensitivity	1	0.395774	1
Specificity	0.942105	0.896084	0.969275
For any particular test result, the probability that it will be:			
Positive	0.07732	0.045437	0.126705
Negative	0.92268	0.873295	0.954563
For any particular positive test result, the probability that it is:			
True Positive	0.266667	0.089136	0.551675
False Positive	0.733333	0.448325	0.910864
For any particular negative test result, the probability that it is:			
True Negative	1	0.973813	1
False Negative	0	0	0.026187
likelihood Ratios: [C] = conventional [W] = weighted by prevalence			
Positive [C]	17.272727	9.733175	30.6526
Negative [C]	0	0	NaN
Positive [W]	0.363636	0.148887	0.888133
Negative [W]	0	0	NaN

Third Phase: Agreement of Estimations of Individual Shoulder Disorders Between the Testers

General agreement over all (7) categories with 2 to 3 (median 3) raters per subject

Cohen's kappa (Landis-Koch extension):

Response	Kappa	se	z(for k=0)	Probability
1	0.772336	*	*	*
2	0.730965	*	*	*
3	0.624838	*	*	*
4	0.655137	*	*	*
5	0.67323	*	*	*
6	0.772131	*	*	*
7	0.527159	*	*	*

Combined (Fleiss-Nee-Landis test):

Kappa = 0.710313 (se = *)
z (for k = 0) = *
* number of ratings per subject not constant, so tests do not apply

Crosstabs

Row variable (first classifier): **A-1**
Column variable (second classifier): **B-1**

	1	2	3	4	5	6	7
1	33	1	0	1	0	0	0
2	5	41	0	4	1	1	0
3	1	0	7	1	1	0	0
4	1	0	0	41	0	2	0
5	0	0	0	1	15	0	0
6	1	0	0	0	1	30	0
7	1	0	0	0	0	0	4

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 88.14%
Expected agreement = 19.13%
Kappa = 0.853397 (se = 0.034009)
95% confidence interval = 0.786741 to 0.920054
z (for k = 0) = 25.093316
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:

1	0.8333	0.6667	0.5	0.3333	0.1667	0
0.8333	1	0.8333	0.6667	0.5	0.3333	0.1667
0.6667	0.8333	1	0.8333	0.6667	0.5	0.3333
0.5	0.6667	0.8333	1	0.8333	0.6667	0.5
0.3333	0.5	0.6667	0.8333	1	0.8333	0.6667
0.1667	0.3333	0.5	0.6667	0.8333	1	0.8333
0	0.1667	0.3333	0.5	0.6667	0.8333	1

Observed agreement = 95.79%
Expected agreement = 65.54%
Kappa = 0.877846 (se = 0.049453)
95% confidence interval for kappa = 0.78092 to 0.974773
z (for kw = 0) = 17.751098
P < 0.0001

Scott's pi
Observed agreement = 88.14%
Expected agreement = 19.25%
Pi = 0.85318

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 12.3539 df = 6 P = 0.0545

Symmetry (generalised McNemar) chi-square = 17.666667 df = 21 P = 0.67

Crosstabs

Row variable (first classifier): **A-1**
Column variable (second classifier): **C-1**

	1	2	3	4	5	6	7
1	28	1	0	2	0	0	0
2	0	37	0	10	3	0	1
3	0	0	15	1	0	0	1
4	1	2	0	30	0	4	4
5	0	0	0	0	12	1	0
6	2	0	0	1	0	32	0
7	0	0	0	0	0	0	6

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 82.47%
Expected agreement = 17.59%
Kappa = 0.787322 (se = 0.032308)
95% confidence interval = 0.724 to 0.850644
z (for k = 0) = 24.369486
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:

1	0.8333	0.6667	0.5	0.3333	0.1667	0
0.8333	1	0.8333	0.6667	0.5	0.3333	0.1667
0.6667	0.8333	1	0.8333	0.6667	0.5	0.3333
0.5	0.6667	0.8333	1	0.8333	0.6667	0.5
0.3333	0.5	0.6667	0.8333	1	0.8333	0.6667
0.1667	0.3333	0.5	0.6667	0.8333	1	0.8333
0	0.1667	0.3333	0.5	0.6667	0.8333	1

Observed agreement = 92.61%
Expected agreement = 64.67%
Kappa = 0.790864 (se = 0.048198)
95% confidence interval for kappa = 0.696399 to 0.88533
z (for kw = 0) = 16.408795
P < 0.0001

Scott's pi
Observed agreement = 82.47%
Expected agreement = 17.71%
Pi = 0.787016

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 13.160963 df = 6 P = 0.0405

Symmetry (generalised McNemar) chi-square = 21.466667 df = 21 P = 0.4308

Crosstabs

Row variable (first classifier): **B-1**
Column variable (second classifier): **C-1**

	1	2	3	4	5	6	7
1	33	1	1	2	1	1	0
2	0	41	0	0	0	0	0
3	0	0	7	0	0	0	0
4	1	0	0	41	0	1	0
5	0	0	1	0	13	1	0
6	1	0	0	0	1	41	0
7	0	0	0	0	0	0	5

General agreement over all categories (2 raters)

Cohen's kappa (unweighted)
Observed agreement = 93.78%
Expected agreement = 19.17%
Kappa = 0.923077 (se = 0.034306)
95% confidence interval = 0.855838 to 0.990316
z (for k = 0) = 26.906949
P < 0.0001

Cohen's kappa (weighted by 1-abs(i-j)/(1-k))
ratings weighted by:

1	0.8333	0.6667	0.5	0.3333	0.1667	0
0.8333	1	0.8333	0.6667	0.5	0.3333	0.1667
0.6667	0.8333	1	0.8333	0.6667	0.5	0.3333
0.5	0.6667	0.8333	1	0.8333	0.6667	0.5
0.3333	0.5	0.6667	0.8333	1	0.8333	0.6667
0.1667	0.3333	0.5	0.6667	0.8333	1	0.8333
0	0.1667	0.3333	0.5	0.6667	0.8333	1

Observed agreement = 97.24%
Expected agreement = 64.07%
Kappa = 0.92309 (se = 0.050582)
95% confidence interval for kappa = 0.823952 to 1.022228
z (for kw = 0) = 18.249539
P < 0.0001

Scott's pi
Observed agreement = 93.78%
Expected agreement = 19.19%
Pi = 0.923063

Disagreement over any category and asymmetry of disagreement (2 raters)

Marginal homogeneity (Maxwell) chi-square = 12 df = 6 P = 0.062

Symmetry (generalised McNemar) chi-square = 5.333333 df = 21 P = 0.9998