DPhil

Age related tendon degeneration: The relationship between rotator cuff tears, shoulder pain, and functional loss.

Hannah Hinsley MA (Cantab) MBBS (AICSM) MRCS (England)

Worcester College & The Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences

Supervisors: Professor Andrew J Carr & Professor Nigel K Arden

Abstract

Background

The shoulder is the third most common site of musculoskeletal symptoms with an estimated 20% of the population reporting symptoms at any given time. Rotator cuff tears are the most common shoulder disorder, and the major debilitation patients seek help for is pain. Full thickness rotator cuff tears have an estimated prevalence of between 7 and 27%, but studies have shown that not all of these are symptomatic. Many case series have been set in shoulder clinics and may have drawn false causality between rotator cuff tears and pain. This study uses a community population-based cohort to determine the epidemiology of rotator cuff tears and test the association between rotator cuff tears, pain, and functional losses.

Methods

463 consecutive subjects (926 shoulders) have undergone a multidisciplinary assessment of their shoulders including high-definition ultrasound, the Oxford shoulder score and myometric strength testing. Individuals were part of the Chingford 1000 women cohort, which is a 20-year-old longitudinal population study comprising 1003 women aged between 64 and 87, and is representative of the population of the UK.

Results

The population prevalence of full-thickness tears was 22.2%, of which 4.6% were bilateral which increased significantly with age. The prevalence was greater in the dominant arm with a 1.64 increase in relative risk. The population prevalence of all tendon abnormalities was 59.3%, of which 30.2% were bilateral, increasing with age.

Although 48.4% of full-thickness tears were asymptomatic there was an association between rotator cuff tears and patient reported symptoms. The relative risk of symptoms compared to normal tendons was 1.79 for abnormal tendons, 2.20 for full-thickness tears<2.5cm, and 4.74 for full-thickness tears >2.5cm ($p<0.001$). Individuals with at least one full-thickness tear had a relative risk of symptoms 1.97 that of those with bilateral normal tendons ($p<0.001$).

Quantitative shoulder strength reduced with age, (10.2-16.2%, $p<0.001$), the non-dominant arm (4.9%, $p<0.001$), and the presence of pain (10.8%, $p<0.001$). Rotator cuff tears had no independent effect, but a significant interaction with age. Strength was preserved in the under 70's irrespective of rotator cuff tear, but in the over 70's there was decrease in strength of between 33% and 39% irrespective of pain ($p=0.004$).

Conclusions

This study provides an epidemiological basis to the understanding of rotator cuff tears. The prevalence of full-thickness tears in a female population aged 64-87 was 22.2%. Although only half of all full-thickness tears are symptomatic there is a statistically significant increase in the likelihood of symptoms with increasing tear pathology. There is also an association with loss of strength but only in the over 70's.

Table of Contents

Acknowledgements

Whilst I alone am responsible for all of the content and work within this thesis, it is the product of years of interaction and inspiration from a significant number of colleagues, whom without the end product would not have been achieved. I would therefore like to thank the following individuals and organisations:

- My supervisors Professor Andrew Carr and Nigel Arden, for their time and patience and continual support over the years of this thesis. Not only was their input key in developing the framework of the thesis, but much time and support devoted to discussions of results and interpretation of data. I would also like to thank their teams at the Nuffield Department of Orthopaedics and Musculoskeletal Sciences who have offered support and help with data management, namely Bridget Watkins, Kim Wheway, and Alison Turner.
- Arthritis Research UK, for their funding and support of the project. In September 2010, at a competitive process I was appointed as an ARUK orthopaedic clinical research fellow, and the project was awarded £190,361.00 to cover costs to completion.
- The Chingford Study and the team devoted to running the year 20 visit at the Silverthorne Medical Centre, Chingford, North London. Firstly I would like to thank Maxine Daniels, organiser and administrator of the year 20 visit, who went above and beyond her duty to enable such a vast amount of data collection, and without whom this thesis would not have been possible. I would like to thank both Tim Spector and Debbie Hart, co-founders of the Chingford study. Further I would like to thank Dr Anushka

Soni, fellow clinical research doctor who was also using the Y20 Chingford study for her support and morale through the process, along with academic discussions on data interpretation.

- Mr Alex Nichols, fellow Orthopaedic SpR, who was responsible for data collection in 143/465 cases seen at the Y20 Chingford study when I was not able to attend in the early stages of the process. With out his contribution to data collection numbers seen in this study would have been significantly reduced. Also Mr Michael Daines, fellow Orthopaedic SpR, who helped to develop the data collection proforma used in the Y20 Chingford visit.
- Gemma Wallace, NDORMS statistician, whose long continual support was key in designing statistical models and advising on statistical techniques. Although the statistics were performed solely by myself Gemma has verified and checked the results, and provided academic discussion around their application through out the entire process
- Mr Jonathan Rees for allowing me to partake in the learning curve study that was used to validate ultrasound skills, and is presented in section 1:1.
- Diana Oag, my mother for proof reading the thesis.

Work from this thesis has been presented at multiple international meetings and are listed below:

- American Academy of Orthopaedic Surgeons (AAOS), San Diego, USA 2011 (Poster)
- International Tendon Symposium, Umea, Sweden, 2010 (Podium)

- European Federation of National Associations of Orthopaedics and Traumatology (EFFORT), Copenhagen, Denmark 2011 (Podium)
- European Society for Surgery of the Shoulder and Elbow (SECEC), Nice, France 2011 (Podium) & Dubrovnik, Croatia 2012 (Podium)
- British Elbow and Shoulder Society (BESS), Torquay, UK 2012 (3 Podium presentations including winner of the Lipman Kessel Best paper prize)
- ARUK Clinical Fellows Research Meeting, Loughborough, UK 2016 (Podium and winner of the best presentation prize)

Publications to date from this thesis include:

Hinsley H, Nicholls A, Daines M, Wallace G, Arden N, Carr A. Classification of rotator cuff tendinopathy using high definition ultrasound. Muscles Ligaments Tendons J. 2014 Jul;4(3):391–7.

Lastly I should like to thank my husband David, and three children Samuel (5), Annabel (3) and Isla (1), who have provided fun and laughter throughout but made it as difficult as possible for me to achieve finishing this thesis!

Introduction

Background information

Musculoskeletal pain is the most common source of disability in the Western world with a UK budget for management of around £10 billion. The shoulder itself is the third most common site of musculoskeletal pain (1), with an estimated 20% of the population reporting pain at any given time (2). Rotator cuff tears are the most common shoulder disorder treated by orthopaedic surgeons. Due to the high disability rates and thus costs (3,4) there is a wealth of literature on the management of rotator cuff tears, and there are many different treatments options, none of which are standardised. One of the possible reasons for this lack of management consensus is the significant variation in clinical manifestations of rotator cuff tears and the lack of understanding of the basic epidemiology and natural history of rotator cuff tears.

The published literature lacks consensus on the population prevalence of full-thickness rotator cuff tears. Wide variations in prevalence have been estimated from cadaveric and radiological studies in both symptomatic and asymptomatic shoulders along with some retrospective cohort studies. However none of these studies have been population based representative of western demographics, and have been subject to significant selection bias. Furthermore, the prevalence of full-thickness tears is not the complete story. To achieve the ultimate goal of understanding the natural history of rotator cuff tears, it is also important to know the epidemiology of tendon degeneration that potentially precedes a full-

thickness tear. Estimations of the prevalence of partial-thickness tears have been speculated in a number of studies, but again none of which are population based. There are no studies to date that have aimed to determine the prevalence of rotator cuff tendinopathy, from the abnormal enthesis up to full-thickness tears using a general population cohort.

Less well established in the existing literature, particularly in relation to the general population, is the relationship between pathology and symptoms. With regards to pain and subjective functional deficits, studies have demonstrated that even within individuals, clinical manifestations of rotator cuff tears can be varied, and attempts have been made to identify factors that may predispose tears to being symptomatic. However to date, no studies have yet explored the associations between rotator cuff tears, and pain or functional loss in a general population cohort. Also no studies have looked either at the relationship of pathology and symptoms with degenerate tendons that are not torn, or the possible change in severity of symptoms across different stages of pathology. This would provide baseline epidemiology to the understanding of the natural history of rotator cuff tears.

Shoulder function can be measured quantitatively as well as qualitatively using strength testing. As this often forms integral parts of shoulder scoring systems it is important to understand how rotator cuff tears and stages of tendon degeneration can affect these, and how they are affected by pathology or pain. This is because not only are the scores used by orthopaedic surgeons to measure responses to treatment but also by general practitioners as cost effective

screening tools to determine who needs treatment. Some studies have attempted to look at strength but only in asymptomatic shoulders. No study to date has measured shoulder strength in association with rotator cuff tear in a general population cohort.

Thesis aims and objectives

The overall aim of this study is to gain a broader understanding of the epidemiology of rotator cuff tears in the general population in the western society both with respect to the prevalence and the association of symptoms. This will provide a foundation to beginning to understand the natural history of rotator cuff tears, which will ultimately have a clinical impact in determining the management and treatment of individuals with shoulder pain and rotator cuff tears.

The specific aims of the study are to:

1. Define the prevalence of different stages of rotator cuff tendinopathy in the general population.
2. Test the hypothesis that rotator cuff tendinopathy is associated with pain and qualitative functional loss.
3. Test the hypothesis that rotator cuff tendinopathy is associated with a quantitative loss in motor strength.
4. Determine whether the severity of qualitative and quantitative symptoms differs in co-ordination with the severity of pathology.

The above aims will be determined using an established general population cohort containing women aged between 63 and 85 years of age, of which details are given in section 1:1.

Section 1: Methods used in the study

1: The general population cohort used: The Chingford Study

The Chingford 1000 women study is an ethically approved well-described prospective population-based longitudinal study of osteoarthritis and osteoporosis. The cohort comprises 1003 white Caucasian women derived from the register of a large general practice in Chingford, North London (5-7). In 1989, 1353 women in the age range 44-67 (mean 54.2) from an age/sex register of over 11000 patients in Chingford were invited to participate in the original study. 1003 were recruited to the study (6 died, 66 had moved away and 278 refused or did not respond). In 1989 the women were representative of women in the UK general population with respect to weight (67 vs. 65 kg), height (162 vs. 161cm), BMI (25.6 vs. 25.4) and smoking characteristics. 98% of the cohort were white, and predominantly middle social class (6).

Since 1989 the cohort has been regularly followed up. It is listed by the NIH as an important epidemiological resource and one of few such cohorts with wide ranging musculoskeletal data. It has been extensively characterised and represents a general population cohort, albeit only women. Over the 20 years clinical, anthropometric, psychological, radiologic and metabolic data has been collected.

This thesis will use data collected at the 20 year follow up of this general population cohort. 516 women attended the year 20 follow-up visits. Of the original 1003, 158 women had died, 111 were unable to attend and 218 had

either moved away, dropped out or were lost to follow up. Of the 516 the median age was 71, and mean BMI was 27.8. A musculoskeletal assessment, including the Oxford shoulder score, and bilateral shoulder ultrasound examination was performed in 464 women. Due to lack of shoulder examiner on particular dates 52 individuals did not undergo this musculoskeletal examination and ultrasound. The median age of the 464 was 71, and mean BMI was 27.8. Of the 464 women, 446 had shoulder strength testing performed. The discrepancy is due to failure of equipment on 2 days of the study. The median age of the 446 was 71, and mean BMI was 27.8. As the participants were selected at random to attend the study, there is no selection bias for the groups examined compared to the 516 who attended the full year-20 visit, which is supported by the demographic details.

The local ethics committee approved the study and consent was obtained from each woman (Outer North East London Research Ethics Committee (formerly Barking and Havering and Waltham Forest RECs), LREC (R&WF) 96).

2: Determining the classification of rotator cuff pathology using high definition ultrasound

1: Background

IMAGING OPTIONS

Imaging plays an important role in the diagnosis and management of the painful shoulder. Ultrasound, MRI and MR arthrography are all used in both the diagnosis of rotator cuff tears, and also the extent of pathology. Studies comparing imaging findings to the gold standard arthroscopic findings have shown equal accuracy between MRI and ultrasound for both full and partial thickness tear detection. In a large systematic review, de Jesus et al found ultrasound sensitivity and specificity rates for the detection of full-thickness tears to be 92.3% and 94.4% respectively compared to 92.1% and 92.9% for MRI. For partial thickness tears these fell to 66.7% and 93.5% for ultrasound and 63.6% and 91.7% for MRI (8-10). In each case MR arthrography was the most accurate but is considerably more invasive. Ultrasound and MRI have also been found to have the same accuracy when quantifying tear sizes: 87% and 80% respectively (11).

Due to the same levels of accuracy ultrasound is considered the most cost-effective imaging modality, as it is cheaper, quicker, and more acceptable to the patient than MRI. It also offers a dynamic examination that can be related to

clinical findings, and can be performed by orthopaedic surgeons in the clinic providing that adequate training has taken place.

HIGH DEFINITION ULTRASOUND VALIDITY

High definition ultrasound has now been extensively validated, and achieves high levels of accuracy for detecting or ruling out full-thickness tears. From systematic reviews of the literature pooled sensitivities and specificities are 0.95 and 0.96 respectively. Additionally the probe used has been investigated and higher sensitivity figures of 0.98 are reported when using a probe of 10MHz or greater (12,13). The major limitation with high definition ultrasound is the ruling out of partial thickness tears. Here sensitivities fall to 0.72 whilst specificities remain high at 0.93 (12).

Detection errors and measurement errors have also been examined (14,15). Misdiagnosis is uncommon with full-thickness tears. However with partial thickness tears it is far more common and the majority of errors are in missed diagnosis. In terms of measurement, errors determined by those estimated outside of 5mm from the arthroscopic measurement, are more common when measuring retraction than width of tear. Errors are reported in 25% of retracted measurements and 12% of width measurements: This though is probably an overestimate of errors due to the sample being composed of predominantly large tears (15).

Analysis of inter-observer variability has also shown to be low when comparing consultant musculoskeletal radiologists. Agreement is found in 96% of cases, and

where there was disagreement this was not major (16). Accuracy levels also depend upon the operator. With the advent of high-resolution portable machines, ultrasound has become available to clinicians to use in clinics, and is not restricted to use by musculoskeletal radiologists. However questions have been raised as to the appropriateness of this as ultrasound remains an operator dependent technique. Numerous studies have demonstrated that experienced orthopaedic surgeons have similar high levels of accuracy when diagnosing full-thickness tears to musculoskeletal radiologists, which is statistically better than general radiologists and sonographers(13,17-20).

In terms of what equates to 'experienced'; the British Society of Shoulder and Elbow Surgeons (BESS) have produced evidence-based guidelines for best shoulder ultrasound practice by orthopaedic surgeons (www.bess.org.uk). This involves the attendance on a shoulder ultrasound-training course, followed by a brief period of familiarity scanning. Following this a training protocol should be adopted, during which shoulders are scanned pre-operatively on the day of surgery and results compared to arthroscopic findings. This should be reviewed after 50 and 100 scans and predictive values compared to published results. Once achieved the clinical use of ultrasound can be recommended (21).

CURRENT CLASSIFICATION SYSTEMS

Tendons are classified into not torn, partial tears and full-thickness tears. However there are a number of surgeon classifications that attempt to subdivide full-thickness tears into smaller subgroups. These include Cofield (1982) and

Bateman (1984) who divided tears into small (<1cm), medium (1-3cm), large (3-5cm), and massive (>5cm). The Southern California Orthopaedic institute derived a similar classification of small (pinhole), moderate (<2cm including only one tendon with no retraction), large (an entire tendon with retraction), and massive (two or more tendons with retraction). Other systems have evolved around shape (Ellman), retraction extent (Patte) or tear position (Habermayer). There is no validated consensus on how to subdivide full-thickness tears according to size or morphology.

A number of studies have attempted to measure the true anatomical footprint of the supra-spinatus tendon in its average insertion width in the A-P plane. Although there is no absolute consensus most musculoskeletal radiology texts quote an A-P width of 25mm (22). Other studies have been performed on cadavers. Ruotolo used callipers to measure 48 cadaveric specimens with a mean age of 72 years, and reported mean widths of 25mm (23), whilst Roh reported mean widths of 21mm +/- 3mm (24). Other studies have commented on the difficulty in making measurements due to the inter-digitation of supra-spinatus and infra-spinatus muscle units and tendon fibres. Minagawa reported that the anterior 12.5mm consisted solely of supra-spinatus fibres, and posteriorly another 10mm was a combination (25). Mochizuki reported isolated supra-spinatus footprints to be as small as 16mm, with the majority of the footprint to be anterior fibres of infra-spinatus sweeping forwards and thus having a combined footprint of 33mm (26).

It has been observed in this unit whilst performing ultrasounds of patients with shoulder pain that there is a subgroup of patients with tendinopathies and pre-tear abnormalities of the tendon enthesis. In these patients the normally smooth appearance of the tendon bone interface is replaced with a craggy uneven interface. This appearance is very easy to determine with ultrasound and could represent a unique un-described subgroup of rotator cuff tendinopathy.

2: Aims and objectives

The overall aims of the thesis are to determine the prevalence of rotator cuff tendinopathy in the general population, and to determine the association of symptoms with pathology. This includes the relationship of severity of pathology and severity of symptoms along with potential differing symptoms between different tear sizes. In order to achieve this a valid classification of rotator cuff tendinopathy must be used, which to date there has not been reported in the literature.

The aims of this preliminary study are therefore to define a valid classification of rotator cuff tendinopathy using high-definition ultrasound that can be used in the out-patient setting, which has discriminate validity between groups and based upon anatomical and clinical principles.

We aim firstly to validate abnormal tendon enthesis as a distinct from normal group as this is easily identified using ultrasound. Secondly, in view of the fact that the literature suggests that we have poor sensitivity at detecting partial

thickness tears, to determine whether these have similar characteristics to the abnormal tendon and should thus be grouped together. Lastly, as the literature supports good specificity with tear size determination, we aim to describe a valid sub-categorization of full-thickness tears. Our hypothesis is that the clinical presentation of isolated single tendon tears will be different to that of larger multi-tendon tears and thus according to the anatomical literature, an ultrasound measurement of a single tendon tear would be up to 25mm on the A-P plane. Measurements will only be taken in the A-P plane due to the superior accuracy that is described above.

3: Patients and methods

SETTING AND SUBJECTS

The study participants were selected from the Chingford Study (as described in section 1.1). With consent and ethical approval a musculoskeletal assessment, and bilateral shoulder ultrasound examination was performed in all of the 465 women who voluntarily attended the 20-year follow up visit. Pain was defined as the presence of shoulder pain in the previous 4 weeks, and severity defined using the validated numeric rating scale (NRS).

ULTRASOUND EXAMINATION PROTOCOL

The procedure protocol was fixed and performed by a single operator (343 by Ms HCL Hinsley (Orthopaedic SpR and author), and 123 by Mr A Nichols (Orthopaedic SpR)). Data was recorded from the left shoulder then the right on a specific data collection sheet.

The ultrasound examination was performed using the GE voluson i portable ultrasound machine with a 10-16MHz linear probe. The protocol for ultrasound examination of the shoulder is laid out below, and is based upon the recommendations of the Nuffield Orthopaedic Centre Musculoskeletal Radiology Department protocol. It is performed from standing behind the patient who is seated on a stool to allow easy movement of the arm.

1. Examination of long head of biceps.

   The hand is placed on the thigh with the palm facing upwards. The biceps tendon is found in the intertubercular groove of the humerus. Comment is made on the presence of the tendon, any fluid surrounding the tendon and the stability of the tendon in the grove when the arm is externally rotated.

2. Examination of subscapularis.

   With the arm maximally externally rotated the subscapularis tendon is viewed in the transverse and longitudinal planes from its insertion on the lesser tuberosity to the point at which it becomes hidden behind the coracoid process. The quality of the tendon and the presence/absence of a full thickness tear were noted.

3. Examination of the supraspinatus tendon.

   The subject's hand is placed on their back with the palm facing the lumbar spine. This extends and internally rotates the shoulder bringing the tendon out from underneath the acromium where it can now be visualized using ultrasound. The tendon is examined anteriorly in a near sagital plane. The tendon is examined in the

longitudinal plane as it inserts into the greater tuberosity, and in the transverse plane from the leading edge backwards. Any fluid in the overlying bursa is commented upon, along with the tendon quality and presence of full-thickness tear. If a full thickness tear is present this is measured on the transverse view, and recorded in cm.

4. Examination of the infraspinatus tendon.

   By placing the subject's hand on the contralateral shoulder into a flexed and adducted position the infraspinatus tendon is found distal to the spine of the scapula in the longitudinal view. Again the quality and presence of a full- thickness tear are noted.

Tendons were classified into one of four working groups, and all tear sizes if present documented. These working groups were created as a result of a consensus meeting, involving Professor Andrew Carr, supervisor and senior shoulder surgeon at the Nuffield Orthopaedic Centre. It was felt that from personal observation that the abnormalities of the tendon enthesis should be included as these represent a clear easily detectable difference on ultrasound to the normal tendon group, and this could potentially provide new evidence for pre-tear conditions. Partial-thickness tears were included as a separate group as they are clinically managed differently to full-thickness tears.

The groups were as follows (see figure 1):

Normal tendon

The tendon has a normal homogenous appearance throughout, and where it inserts onto the greater tuberosity there is no abnormal reaction-taking place on the bone side.

Abnormal tendon

The tendon may appear abnormal with loss of the homogenous appearance or appear abnormally thin. There is an abnormal tendon enthesis, which is clearly visualised as a ragged greater tuberosity. Fluid may be present in the bursa.

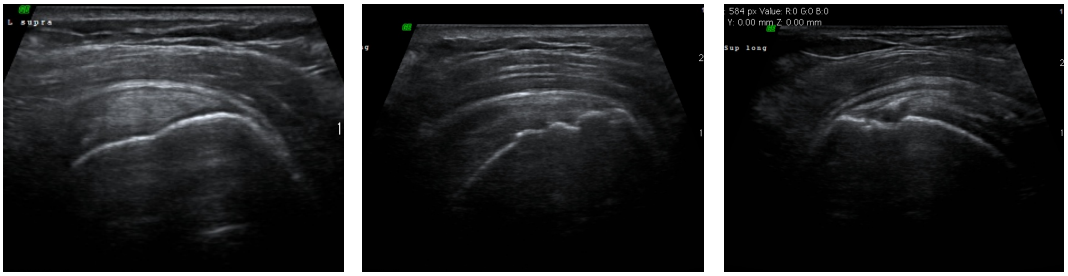
Partial-thickness tear

A lucent patch is visualised in the tendon either on the articular or bursal surface, but there are continuous tendon fibres inserting into the greater tuberosity. In all cases an abnormal tendon enthesis was observed (even in those with bursal sided tears). Fluid is often seen in the bursa.

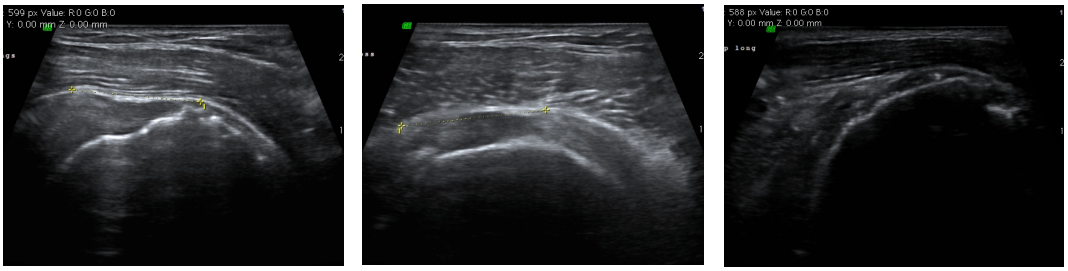
Full-thickness tear

A lucent patch is visualised running through the full-thickness of the tendon. The width of this in the sagittal plane is the tear size. This is associated with a concavity in the bursal surface, which is either directly visualised or becomes apparent with pressure from the probe. The tendon enthesis is abnormal and there is fluid in the bursa. In larger tears no tendon is often visible. The surface of the greater tuberosity is abnormal and a heterogeneous substance fills the space left by the tendon. In smaller tears a tendon edge may be

visible in the longitudinal plane and this measurement is the degree of retraction.



Normal tendon | Abnormal tendon | Partial-thickness tear



Full-thickness tear <2.5cm (longitudinal view) | Full-thickness tear <2.5cm (transverse view) | Full-thickness tear >2.5cm (Longitudinal view)

Figure 1: Rotator cuff classification

Training and validation

Training, followed the protocol laid out in the BESS guidelines (www.bess.org.uk). The initial training took place on the *Skills Course in Musculoskeletal Ultrasound* held in Oxford in January 2010. Further familiarity experience took place in the outpatient department at the Nuffield Orthopaedic Centre, under the supervision of professor Andrew Carr, who is trained and well-established in shoulder ultrasound.

Sonogropher validation took place using a local learning curve study (21). In this study, 50 pre-operative patients awaiting shoulder arthroscopy underwent an ultrasound scan of their shoulder on the morning of their surgery. Participants were blinded to the diagnosis of the patients being scanned. The results were compared to the same day arthroscopy findings.

Two examiners performed the ultrasound examinations in this study, both of whom underwent the same training protocols. Inter-observer validity was performed using statistical comparisons between the distributions of diagnosis after adjusting for age and pain. Intra-observer validity, by repeat scans, was only possible for 18 shoulders due to lack of consultation space. Verbal and clinical examinations were not re-performed and no notes from the previous examination were available. The protocol for the ultrasound examination was unchanged.

Statistical analysis

Inter-observer validity was demonstrated comparing the distributions of tendinopathy groups diagnosed on ultrasound after adjustment for age and BMI using the Mann-Witney-U test. Intra-observer validity was demonstrated using a weighted kappa test.

Age, BMI, and pain presence were compared between tendinopathy groups. Age was not normally distributed and BMI was normally distributed. The Mann-Witney-U test, Students T-Test, and Fischer's exact tests were used for non-normal, normal and categorical data respectively. Distributions of tear sizes

across symptomatic and asymptomatic shoulders were compared using the Mann-Witney-U test. Correlations between tear sizes and the NRS were demonstrated using a smoothed best-fit curve and spearman's correlation co-efficient. To determine the optimum tear size as a predictor of pain a sensitivity specificity analysis was performed and a receiver operator curve used to predict the best fit model. All statistics were performed using IBM SPSS Statistics version 22 (IBM, Armonk, NY, USA).

4: Results

1: Training and validation results

The results of the learning curve study are shown below (table 1) (21). The author's accuracy of diagnosing full-thickness tears was 96.1% (weighted kappa 0.959 (p=<0.001)), with sensitivity 94.7%, specificity 96.8%, Positive predictive value 95% and negative predictive value 96.1%. The accuracy of diagnosing partial thickness tears was 86.3% (weighted kappa 0.557 (p=<0.001)), with sensitivity 50%, specificity 95.1%, Positive predictive value 71.4% and negative predictive value 89.1%.

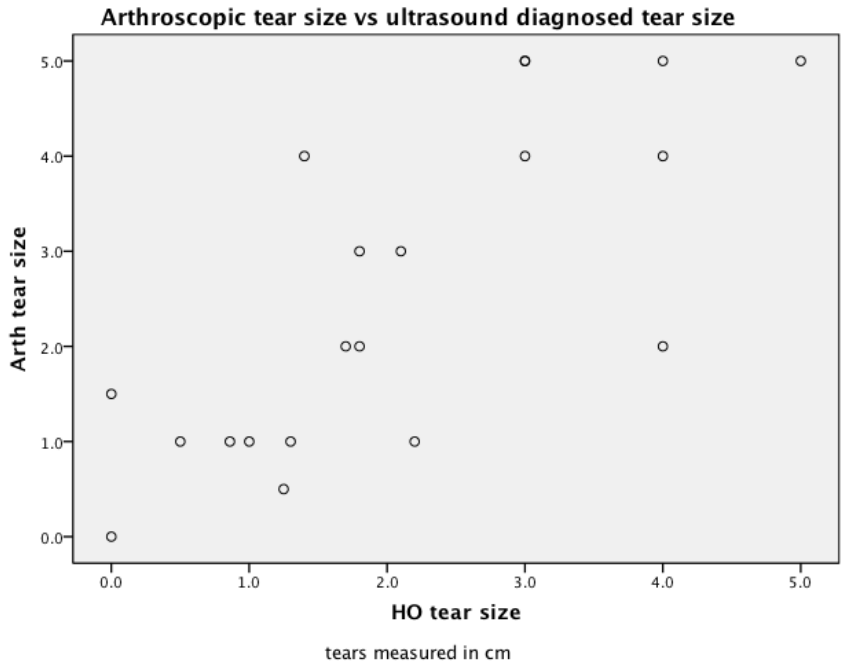
Comparison of ultrasound and arthroscopic diagnosis from the learning curve study

Ultrasound diagnosis	Arthroscopic diagnosis			Total
	Intact	Partial tear	Full-thickness tear	
Intact	20	4	1	25
Partial tear	2	5	0	7
Full-thickness tear	0	1	18	19
Total	22	10	19	51

Table 1: Results of the learning curve study; Comparison of ultrasound and arthroscopic findings

Of the one case in which a full-thickness tear was missed, it was called as an abnormal tendon, which was in fact was an un-displaced full-thickness tear. Of the 5 false negative cases with the partial tears, 4 of the tendons were called abnormal, and one called a FTT. Of the two false positives no tears were present. No comment is made on the footprint of the tendon.

Pearson's correlation coefficient of tear size measurement was 0.761 (95%CI 0.469-0.903, p<0.001)(see graph 1), and when divided into group's accuracy was 85%.

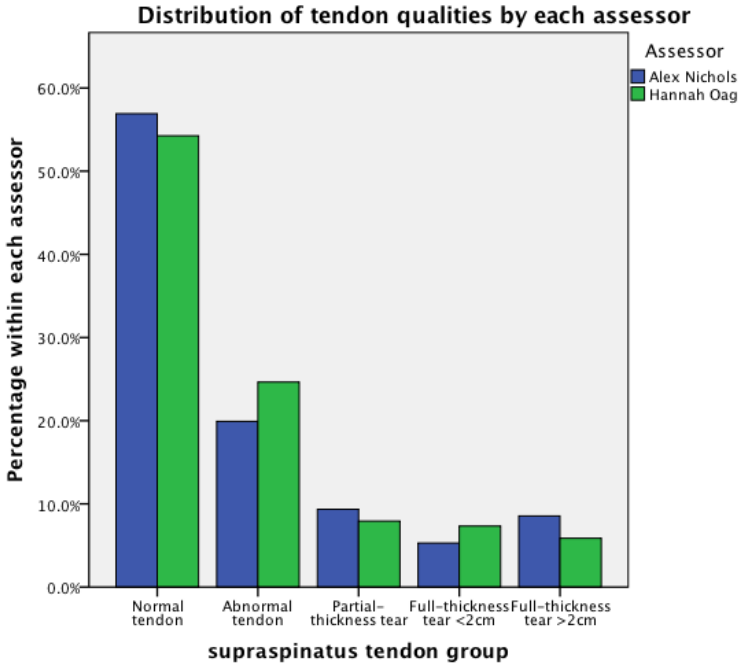


Graph 1: Correlation between tear sizes measured on ultrasound to those reported at arthroscopy

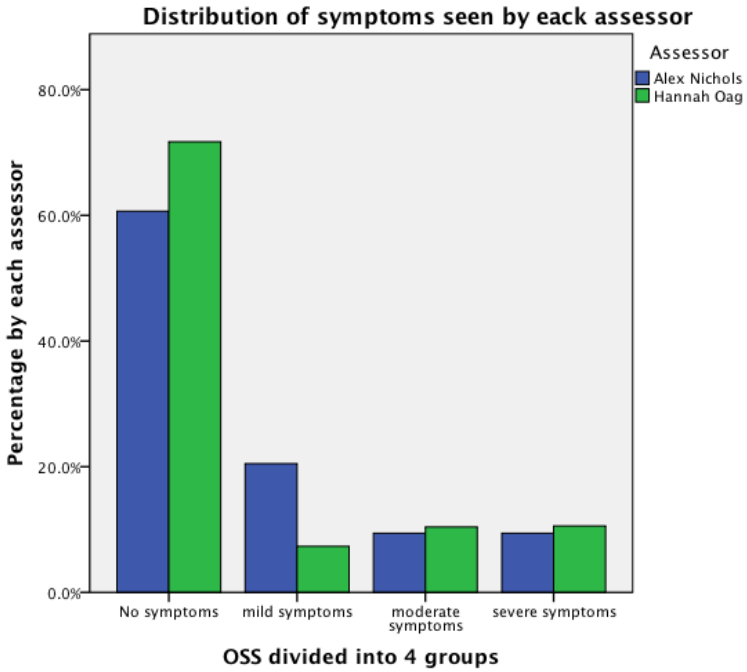
Inter-observer validity within the general population cohort

There was no significant difference, after adjustment for age and pain, (Mann-Witney-U p=0.80) in the distributions of diagnosis scanned by each examiner

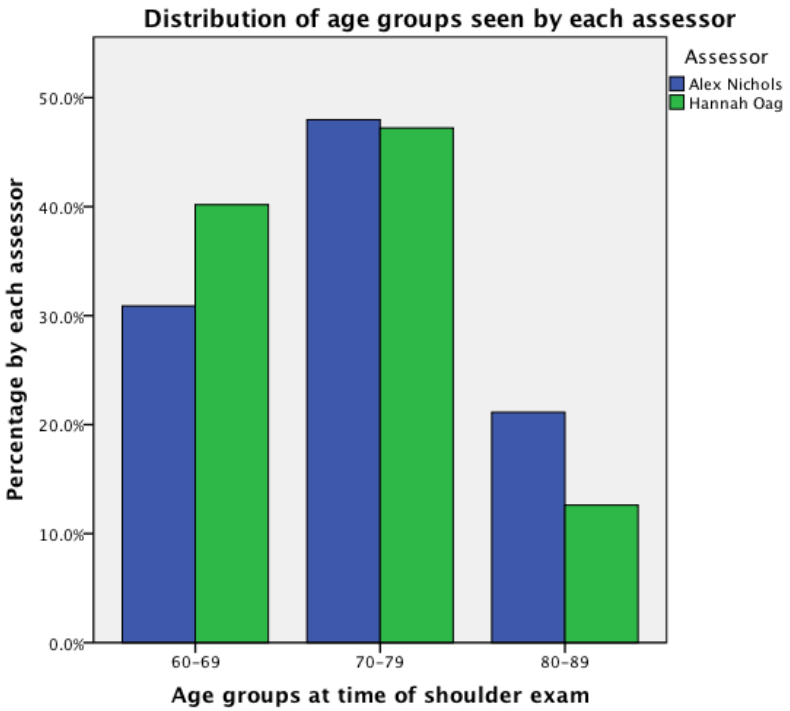
across the cohort. Both examiners scanned two cases and there was 100% agreement within the classification system (see graphs 2-4).



Graph 2: Showing no difference in the distribution of tendon qualities reported by each assessor



Graph 3: Showing no difference in the distribution of symptoms seen by each assessor



Graph 4: Showing no difference in the distribution of ages seen by each assessor

Intra-observer validity within the general population cohort

Of the 18 shoulders re-scanned there were 3 disagreements (see table 2). In two cases a previously diagnosed partial-thickness tear was classified as an abnormal tendon. The other disagreement occurred in the size of the full-thickness tear. Here the original diagnosis was a 1.3 cm full-thickness tear, which was documented as having only a very thin abnormal remnant of tendon visible aside of the tear. Two months later this is recoded as a massive tear with no visible tendon. In order to statistically obtain a measure of agreement the results the abnormal tendon group and partial tear group were lumped together (weighted kappa 0.915 $p<0.001$).

Cross tabulation of original and validated supra-spinatus groups
All shoulders

Original supra-spinatus diagnosis	Validation supra-spinatus diagnosis					Total
	1	2	3	4	5	
Normal tendon (1)	2	0	0	0	0	2
Abnormal tendon (2)	0	7	0	0	0	7
Partial thickness tear (3)	0	2	0	0	0	2
Full-thickness tear <2cm (4)	0	0	0	4	1	5
Full-thickness tear>2cm (5)	0	0	0	0	2	2
Total	2	9	0	4	3	18

Table 2: Comparison between the initial ultrasound diagnosis and that made at a repeat scan. Only 2 cases differed between scans.

2: STUDY RESULTS

926 shoulders were included in the study; of which the clinical, demographic and radiological characteristics are shown in table 3.

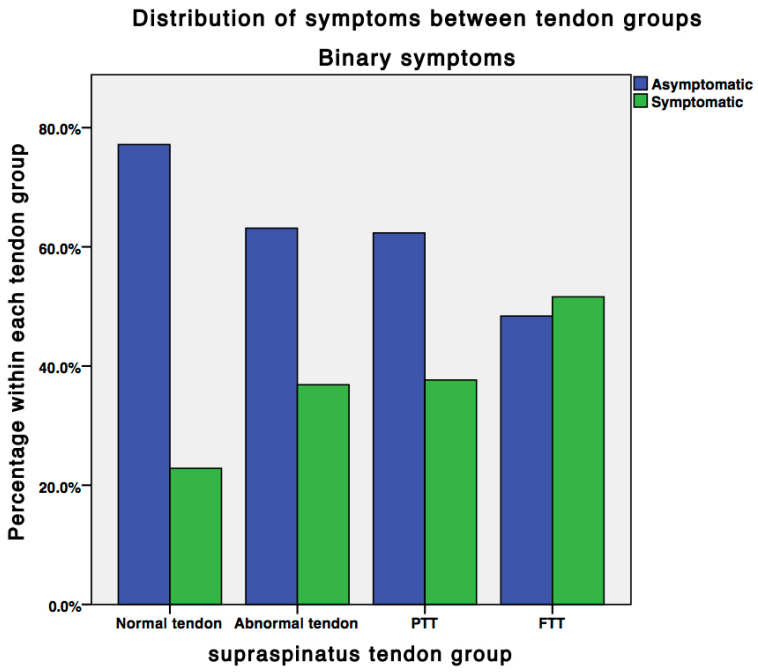
Demographics of shoulders in the study

	N (%)	Median age	Mean BMI	% Pain
Normal	510 (55.1)	71	27.5	22.8
Abnormal	217 (23.4)	73	28.1	36.9
Partial tears	77 (8.3)	74	27.5	37.7
Full-thickness tears	124 (13.3)	74	28.5	51.6
All	926	71	27.8	31.2

Table 3: clinical, demographic and radiological characteristics of the 926 scanned shoulders compared by tendon group.

Abnormal footprints and partial-thickness tears

There was no significant difference in age, BMI or pain between abnormal tendons and partial thickness tears. However age was significantly different when each group was compared to normal tendons (Mann-Witney U $p<0.001$). The proportion with pain was also significantly different between normal tendons and both abnormal tendons (Fishers exact test $p<0.001$) and partial-tears (Fishers exact test $p=0.050$). Both the abnormal tendons and partial tears had a significantly greater proportion of individuals with pain compared to normal tendons, yet no significant difference between each other (see graph 5). Logistic regression models predicted the relative risk of having pain compared to normal tendons was 2.0 ($p<0.001$) for abnormal tendons, 2.1 ($p=0.005$) for partial tears, and 3.6 ($p<0.001$) for full-thickness tears.

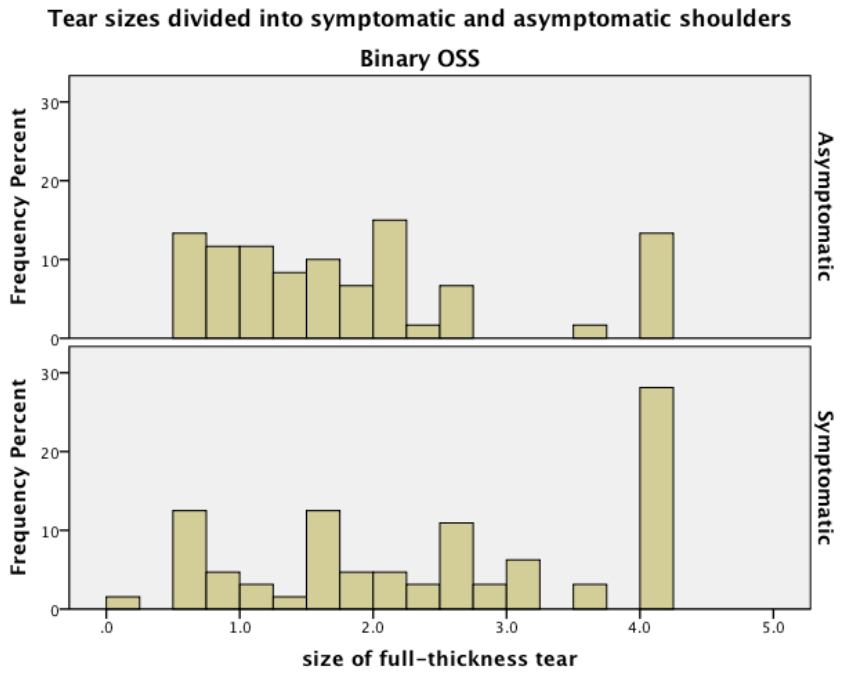


Graph 5: Showing that abnormal and PTTs had a greater likelihood of symptoms then normal tendons but less then FTTs, though there was no difference in symptom likelihood between abnormal and PTTs.

Full-thickness tear sizes

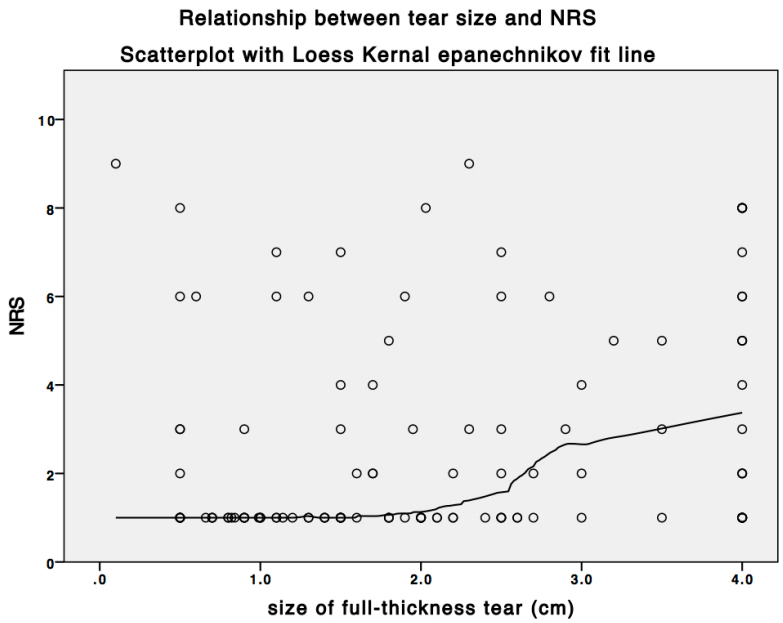
123 full-thickness tears were included in the analysis. There was no difference in BMI between any of the tendinopathy groups. Those with full-thickness tears were significantly older than those with normal tendons (Mann-Witney U, $p<0.001$), but not those with abnormal/partial tears. The proportion of full-thickness tears with pain was significantly different from normal (Fishers exact test $p<0.001$), and abnormal/partial tears (Fisher exact test test $p=0.006$). Logistic regression demonstrated RR of pain compared to normal tendons of 3.6 ($p<0.001$) compared to 2.0 ($p<0.001$) for abnormal/partial torn tendons.

Tear sizes were not uniformly distributed between symptomatic and asymptomatic shoulders, with the median tear size being greater in the symptomatic shoulders (Mann-Witney-U, $p=0.006$)(graph 6).



Graph 6: Comparison of tear sizes between symptomatic and asymptomatic shoulders showing a greater median tear size in the symptomatic shoulders.

When correlating the size of tear with the NRS score, the smoothed best-fit line demonstrates a change in pain profile between tear sizes 2 and 3cm (spearman correlation co-efficient 0.245, p=0.006)(Graph 7).



Graph 7: Relationship between tear size and the NRS, showing a change in NRS scores between 2 and 3 cm tears

A binary logistic regression model predicted that tears would become symptomatic once greater than 2cm, however, firstly this assumes a linear relationship which is not the case, and secondly it is only telling us that the point at which symptoms become more than 50% likely occurs at >2cm.

For tear size to be a predictor of pain, a receiver operating characteristic (ROC) curve was used to define the cut-off values predicting the presence of pain with optimal sensitivity and specificity, and the highest positive predictive value. This occurred between 0-2.5cm and >2.5cm. The sensitivity- specificity analysis is shown in tables 4-6.

Sensitivity – Specificity Analysis

		Condition (Symptom presence)	
		Symptomatic	Asymptomatic
Test (Tear size cut point)	Large tear	True positive A	False positive B
	Small tear	False negative C	True negative D
		Sensitivity = A/(A+C) PPV = A/(A+B)	Specificity = D/(D+B) NPV = D/(D+C)

Table 4: Variables used in the specificity sensitivity analysis.

Sensitivity specificity analysis of different tear size cut points

Groups		Sensitivity (sens)	Specificity (spec)	sens/(1-spec)
Small	Large			
0-2.0	>2.0	0.571	0.700	1.90
0-2.1	>2.1	0.563	0.733	2.11
0-2.2	>2.2	0.547	0.767	2.34
0-2.3	>2.3	0.516	0.767	2.21
0-2.4	>2.4	0.516	0.783	2.38
0-2.5	>2.5	0.453	0.833	2.71
0-2.6	>2.6	0.422	0.833	2.54
0-2.7	>2.7	0.406	0.850	2.71
0-2.8	>2.8	0.391	0.850	2.61
0-2.9	>2.9	0.375	0.850	2.5

Table 5: Demonstrates the closest to perfect test score (sens/(1-spec)) occurred when dividing tear sizes into groups 0-2.5cm and >2.5cm.

Predictive value analysis of different tear size cut points

Groups		PPV	NPV
Small	Large		
0-2.0	>2.0	0.667	0.609
0-2.1	>2.1	0.692	0.611
0-2.2	>2.2	0.714	0.613
0-2.3	>2.3	0.702	0.597
0-2.4	>2.4	0.717	0.603
0-2.5	>2.5	0.744	0.588
0-2.6	>2.6	0.730	0.575
0-2.7	>2.7	0.743	0.573
0-2.8	>2.8	0.735	0.567
0-2.9	>2.9	0.727	0.560

Table 6: The best positive predictive value to determining whether tears were symptomatic also occurred when tear sizes were divided into groups 0-2.5cm and >2.5cm.

5: Discussion

The overall aim of the study was to produce a valid discriminate classification system of rotator cuff tears using high-definition ultrasound. It must aid the overall aims of the thesis in determining firstly the prevalence of different stages of rotator cuff tear in the general population, and secondly whether different stages are associated with differing amounts of symptoms. It also must be able to be applied clinically in an outpatient setting for general use. Therefore ultrasound has to be able to validly determine the difference between groups.

The existing literature reports that high levels of accuracy have been achieved by orthopaedic surgeons in detecting full-thickness tears, but not partial thickness tears. Using the learning curve study we have demonstrated equally comparable levels of accuracy in detecting full-thickness tears but likewise inadequate accuracy in detecting partial thickness tears, where sensitivities fall significantly lower. This evidence would support any classification system that could distinguish between full-thickness tears vs. no full-thickness tears. It would not though, support a separate category for partial thickness tears. From the minimal data that is present regarding the missed partial-thickness tear in the learning curve study, each of these was documented as having an abnormal tendon enthesis. This would support the rationale for having a group in which the tendon is not normal. The personal experience of this department is that although partial thickness tears are difficult to detect, this abnormal tendon enthesis is easy to see on ultrasound. Further to this the Chingford cohort study has demonstrated that pain proportions and demographic profiles for abnormal

tendons and partial tears groups are the same, yet significantly different to normal tendons and full-thickness tears. Thus an abnormal tendon enthesis and partial-thickness tear group combined is a valid and easily identified group, which has not previously been described in the literature.

Along with high levels of accuracy in detecting full-thickness tears the literature also suggests that accuracy can be achieved in measuring tear sizes in the A-P plane. Although at first glance we were unable to achieve the same accuracy as reported in the literature in measuring tear sizes two factors should be considered. Firstly in the learning curve study the 'gold standard' measurement taken during the arthroscopy was taken from the operative notes. None of these had been accurately measured and were given as integers of the nearest centimetre using the naked eye. Secondly if the measurements were divided into two groups 85% accuracy was in fact achieved. Also the overall correlation coefficient was good at 0.761 (95% CI 0.469-0.903, $p<0.001$).

Current classification systems have often divided full-thickness tears into groups. However these have not necessarily been evidence based. Our hypothesis is that single and multi tendon tears would present clinically differently. Thus, based upon anatomical studies we are hypothesising that tears up to 2.5 cm would present potentially with less pain than larger ones. Using the Chingford general population cohort, we have shown that the likelihood of having pain is not uniform across the whole full-thickness tear group, and neither does it change in a linear fashion. We used a sensitivity specificity model to try to determine the point at which full-thickness tears become increasingly

likely to be symptomatic, and defined the cut point using the closest to perfect predictive model. This occurred between sizes 0-2.5cm and >2.5cm. In line with radiological and anatomical studies looking at the supraspinatus footprint this confirmed our hypothesis that the change occurred between single and multi-tendon tears, and likely reflects the altered biomechanics of the shoulder with large multi-tendon tears.

When performing the sensitivity-specificity analysis to determine the two tear size groups we chose the point at which the model was closest to perfect rather than highest sensitivities or specificities. For this particular experiment we have used the specific tear size cut point as the test to predict whether or not symptoms will be present. A test with high sensitivity would place the cut point at a smaller tear size and consequentially more individuals would be predicted to have symptoms. As a result the number of false negatives (small tears with symptoms) would fall; however the number of false positives (asymptomatic tears) would increase, resulting in more asymptomatic tears being predicted as being symptomatic, which is then reflected in a poorer test specificity. In this cohort we must remember though that 22.8% of normal tendons, were in fact symptomatic, i.e. there is a high baseline level of pain, so driving the sensitivity up may not be appropriate as we would be picking up more baseline pain rather than that specific to the tears. A test with a high specificity would place the cut point at a larger tear size and mean fewer individuals were predicted to have symptoms. As a result the number of false positives (large asymptomatic tears) would decrease; however the number of false negatives (small symptomatic tears) would increase and more symptomatic small tears would be predicted to

be asymptomatic which is reflected in poorer test sensitivity. However, with this cohort 48.4% of all full-thickness tears were asymptomatic, and indeed only 77.2% of normal tendons were in fact asymptomatic. Thus driving the specificity upwards would mean missing many symptomatic smaller tears. With respect to the clinical application of the test, we are aiming to determine what tear size is associated with an increased risk of having symptoms. Therefore a balance is required and choosing the closest to a perfect test is represented by sensitivity / (1- specificity). This value also coincided with the highest positive predictive value of pain.

STUDY LIMITATIONS

There were potential limitations in the study methodology. Firstly, using a single clinician to assess symptoms and perform the ultrasound examination. After performing the history and examination the examiner had already been exposed to the primary outcome measure and was thus not blinded. With the hypothesis that rotator cuff tears are associated with pain and functional loss, and that the severity of symptoms is associated with the severity of pathology then knowing symptoms prior to the ultrasound is more likely to produce a positive association between the two. This order and knowledge of symptoms may also bias the prevalence data produced. If pain were in fact more prevalent than pathology, the examiner would have a tendency to report a higher prevalence of pathology. If pain were less prevalent than pathology, then prevalence rates may be under-reported. This could have been avoided by using two independent examiners. This however was not possible due to resources. A simple way to overcome this potential bias would have been to have 2 independent examiners,

where the examiner performing the ultrasound was blinded to the symptom assessment: This however, was not possible with the resources available. To attempt to look at the impact of this bias, a small intra-observer reproducibility study was performed and although demonstrated good reproducibility it ideally needed to be larger. If the protocol had been performed in reverse order the prevalence rates my have been less subject to bias, however there would, I feel, have been a greater the potential for bias to influence the association of symptoms. This is because symptoms are subjective measures recorded by the patients and could have been influenced by the assessing doctor.

Secondly, bias may potentially have been introduced in the shoulder ultrasound protocol. The left shoulder ultrasound was always performed prior to the right side. Thus the right scan was subjected to bias as a result of the left side. In particular with equivocal decisions it is likely that agreement with the opposite shoulder would be recorded. In terms of prevalence, if pathology is less common in the non-dominant arm (usually left thus likely to be scanned first) then the prevalence in the dominant arm could be underestimated. If reversed the opposite would have been true. If independent examiners scanned the shoulders this bias would be reduced; however, inter-observer bias would be introduced. The advantage of always scanning one shoulder first does though reduce data entry errors, and minimise the risk of mixing up left and right.

There is also a potential methodological limitation when using the same cohort to determine a classification that will later be used to determine the primary outcome. Here the Chingford cohort has been used to determine two groups of

rotator cuff tear, which we will then later assess to determine different symptom profiles. We are potentially more likely to reject our null-hypothesis that there is no difference between the groups, by determining the groups based upon the results. It also means that any differences detected between the groups are only applicable to the cohort being studied and may not hold true for other cohorts. Ideally this cut point should have been determined using either pre-existing evidence from the literature or by data collected from another cohort, which was not possible. However the results support our initial hypothesis that single and multi-tendon tears behave differently, and this size was taken from previously performed radiological and anatomical studies. Alternatively samples from the cohort could have been used and then applied to the whole cohort, however with only 128 full-thickness tears numbers were not large enough. In order to validate this cut point it needs to be applied to other cohorts, which will be planned as future work.

6: Summary

Using a general population cohort this study has demonstrated a valid classification of rotator cuff tears using high definition ultrasound that can be applied in the outpatient setting and is based upon anatomical principles and clinical presentations. Each group can validly be identified, and has different profiles in terms of demographics and pain.

The classification system is as follows:

1. Normal Tendons
2. Abnormal enthesis/ partial-thickness tear
3. Single tendon full-thickness tears (0-2.5cm)

4. Multi-tendon full-thickness tears (>2.5cm)

Our justification for this classification is that we cannot accurately detect partial tears yet we can abnormal tendons, and these present clinically the same. There is not a linear relationship between pain likelihood with tear size and this changes at the average supraspinatus width representing single and multi-tendon tears.

This classification system will therefore be adopted through this thesis study.

3: Determining the primary outcome measure: Scoring of shoulder symptoms

1: Background

1: Primary outcome measures required for this study

This thesis aims to determine whether there is an association between rotator cuff tear, pain and functional loss in the general population. The hypothesis states that with increasing tear stage severity symptoms will become firstly more prevalent and secondly more intense.

The specific questions are:

1. Are different stages of rotator cuff tear associated with pain and functional loss?
2. If so, does the severity of pathology correlate to the severity of symptoms perceived?

Rotator cuff tears are biologically thought to be a chronic condition. Thus accurate symptom assessment needs to capture the extent that this affects the patient on a day-to-day basis rather than at one particular given point. As symptoms will fluctuate with different activities and at different times of the day, any assessment would beneficially encompass all of this to reduce the introduction of bias that a one-dimensional, single time point question may bring. In terms of defining symptoms, it must be clear as to whether this is pain

or functional loss or indeed both. Pain and functional loss tend to be intrinsically entwined, for example function may be lost due to pain, however in chronic disease states pain can subside and a functional deficit still remain. Therefore pain and functional loss must be evaluated both together and independently. For this study we therefore need outcome measures to assess pain and functional loss both independently and combined, along with capturing symptoms that may fluctuate.

The scoring system must have been validated to be practical, reliable and valid. Practical means that it must be straightforward to use and appropriate to the cohort being studied. Reliability is a term that reflects the extent to which a measurement gives consistent results on repeat applications and to different cohorts. The validity of a measurement is the approximate truth that the conclusion drawn from the measurement accurately reflects the construct it was intended to. Finally reproducibility reflects the extent to which the measurement is consistent when used by different individuals. In order to be valid for this study the chosen score must demonstrate:

1. Content validity

   The content validity of any test is the systematic examination of the test content to determine whether it covers a representative sample of the domain being investigated. Thus for this study it must:

   a. Measure all the important issues relating to shoulder pain and functional loss.
   b. Target the population in question (females aged 63-85)

c. Measure recent trend in pain/ symptoms rather than at present moment (as a chronic condition)

2. Construct validity

   This is the broad overriding term of validity that refers to the extent to which the operalizations of the construct actually measure what they are theoretically supposed to. This term also includes criterion validity such as convergent and divergent validity. These are the extent to which the measure correlates to other measures that it theoretically should or shouldn't. Thus, for this study it must:

   a. Measure shoulder pain or function and not be subjective to noise from surrounding areas.
   b. Have been validated to study rotator cuff disease
   c. Measure patients' perspectives of the problem.
   d. Demonstrate convergent and divergent validity to other tests.

3. Discriminate validity

   This is the ability an operalization has to distinguish differences between groups, thus, it must:

   a. Be able to distinguish between asymptomatic and symptomatic individuals.
   b. Discriminate between those who believe their symptoms are mild moderate or severe.

2: SHOULDER PAIN AND FUNCTION OUTCOME MEASURES

Codman introduced the concept of the 'end result' in orthopaedics in the 1820's (27). This day, the pursuit of 'best clinical practice', along with health economic planning and patient awareness, has driven the emergence of many scoring systems (28). Scores can be generic, such as the SF36, or shoulder specific, which can be clinician based, such as the Constant-Murley score (29), patient based such as the Oxford shoulder score (OSS) (30,31), or a combination of both such as the American Shoulder and Elbow Score (ASES) (32). When looking at shoulder problems the shoulder specific questionnaires have proved more valid than generic scores (33).

Clinician based assessments allow clinicians to examine the shoulder and also may prevent patient misinterpretation of questions. However, patient based outcome measures (PROMs) are becoming increasingly popular. Firstly, they are less susceptible to bias (34). Secondly, patients and surgeons often will differ in their respective concerns (30,31) and giving prominence to the patient's perspective is preferable as it captures the importance of their own health care objectives, and it has been shown that patients can provide valid and reliable judgements regarding their own health (35,36). Also when compared, shoulder scores have shown week correlations with measured shoulder movements (37). Thirdly, PROMS have greater stability with time compared to clinician assessments (33). And lastly, they can be completed remotely and be returned by post if required (38). All of the scoring systems in common use have been extensively validated to be practical, reliable, valid and responsive.

Shoulder specific PROMs

The OSS is the only PROM that has been developed solely from patient interviews (30,31,33). It has been validated to be used in rotator cuff pathology (39,40) and is valid as a postal questionnaire (38). Its design, however intends its use for detecting change from surgical procedures. It has not been validated in a general population or for use in isolation without repeat measures. In fact one study has shown that the asymptomatic shoulders can have a wide range of scores (41).

Both the Shoulder disability questionnaire (SDQ) and the Shoulder pain and disability index (SPADI) have been validated in the primary care setting (37,42,43). The SDQ is designed using binary answers, to be quick and easy to complete, and cover a full functional limitation profile. It asks only of limitations on the day of survey completion. Due to its binary method of answering it is unable to quantify the severity of individual parts, however, it has highlighted that for patients with shoulder pain the most prevalent problem was pain interfering with sleep. It has also demonstrated that for all individuals with shoulder pain those that sought medical treatment form a general practitioner had a worse score than those who had not (42). SPADI asks about symptoms in the past week and has a 1-10 rating scale for each question. Its pain and disability parts can be used separately or combined. Although it performs well in validation studies there are no questions regarding night pain and only 2 assets of a functional limitation profile are covered (37).

Pain specific PROMs

There are many other scoring systems in existence that focus generically on pain, rather than site-specific problems. These include one-dimensional pain scales such as the visual analogue scale (VAS), the numeric rating scale (NRS) and the verbal rating scale (VRS), all of which only measure pain intensity. Multidimensional scales such as the McGill Pain Questionnaire and the Brief Pain Inventory have been developed to explore the complexity of pain, whilst many others exist exploring specific aspects of pain or psychosocial contributions. Often though for research purposes the one-dimensional scales are less susceptible to bias when comparing pain in different groups (44).

The VAS, NRS and VRS have been all been shown to be valid reliable and appropriate for use in clinical practice (45,46). Although the VAS has proven the most reliable it had higher failure rates in more elderly groups and poorer reliability in those with mild cognitive impairment. For these groups the NRS proved reliable. By nature the VAS and NRS are the most sensitive tools (47).

Pain data can also be collected retrospectively. However, a number of studies have shown that prior musculoskeletal symptoms are poorly recalled. One study followed up patients over a 6-year period and demonstrated that 72% of patients did not recall any pain 6 years later, of these this rose to almost 100% in those with no current pain. Over recalling of symptoms was reported in 37% of patients who had current pain (48). Other studies have supported this concern, and demonstrated that patients find it difficult to articulate how they derived old pain ratings, and there was little consistency across patients (49).

3: Outcome measures used to collect data in the Chingford Y20 study

The Oxford Shoulder Score (OSS)

The Oxford Shoulder Score (OSS) (30) is a validated 12 item PROM that was introduced around 15 years ago, primarily for the assessment of outcomes after shoulder surgery, however its uptake and usage has been extended over time. It has been validated to be reliable, valid and responsive. The content of the OSS is unique as it was developed out of interviews with patients and thus reflects their perspectives. It has been validated and shown to have high internal consistency and good reproducibility. The content has been validated against the constant score and the relevant sections of the Stanford health assessment questionnaire (HAQ) and the SF36 form, which have demonstrated good construct validity. It has shown high sensitivity to change with time. Although its prime objective was to be responsive to change, which it achieves, the reliability of the score also adds some justification for the score to be used in isolation providing it is valid for what is being assessed. It however has not to date been validated in a general population cohort.

The OSS is designed to be joint specific, and by the use of hypothetical questions, aims to be influenced as little as possible by co-morbidities: For example, "due to you shoulder 'could' you do the household shopping on your own?" It has also been validated for use specifically in rotator cuff disease (40).

The OSS contains 12 questions, all regarding the shoulder in the last 4 weeks. Four of which are denoted to pain, and eight to function. Each question is scored from 0-4, with 0 being the worst possible outcome. This then leaves a range from

0-48 in total, with 48 representing the best possible outcome. Although questions can be distinctly divided into pain and function, its use in this way has not been validated.

Application of the OSS to this study

1. Content validity

The OSS allows the assessment of pain and activities of daily living in one score. It has distinct pain and function questions, which could be used independently (and have been done so in various studies), though this separation has not been validated to date.

In terms of pain the 4 main areas that concern patients are targeted; worst pain, usual pain, interference with work and night pain. This recognises that pain is not a single dimensional component, and allows individuals to describe pain in more detail. Validation studies of the SDQ demonstrated that night pain was the most prevalent symptom with shoulder problems, which is highlighted in the OSS. Bias may be introduced if individuals were only commenting on pain during the day.

Symptoms are measured over the last 4 weeks, which biologically fits with our model hypothesis; this is a chronic condition and thus bias could be introduced if individuals are having a good day, which in turn would weaken any associations found.

It also targets the age group in question appropriately. Many criticisms of scoring systems are that the questions are not relevant to the population being examined. Here the activities of daily living that are asked about seem to be appropriate for our population. In younger age groups the questions may not be sensitive enough to detect problems. However this was not a concern in the design of the OSS as the importance was the ability to detect change.

2. Construct validity

The construct validity is good. The OSS has been validated against other scoring systems, including the Constant score, SHQ, SF-36 (33) and the Western Ontario Rotator Cuff score WORC. It has also been validated to study specifically rotator cuff pathology (40).

It is based solely on patient perspectives. Firstly it was developed out of patient interviews and thus intends to capture which issues are important to the individuals completing the questionnaire. Secondly as a PROM, it is not subject to potential clinician bias that is the case in many scores such as the Constant and ASES scores.

It is designed to be free of noise by asking questions hypothetically (30). However, it has been speculated that these hypothetical questions can be misinterpreted when the questionnaire is applied to the general population. Thus if individuals are reporting concurrent symptoms, scores may be worse than expected (41). If this were to affect the more advanced tear stage groups more than the normal tendon group, which seems logical as these are more likely

to have concurrent pathology, then there would be a false increase in the association of symptoms.

3. Discriminate validity

The discrimination validity is the weakest component when using this score for this study's purpose. Although a numeric range is produced and individuals can grade each symptom, the score was not designed to categorise individuals into categories of altering degrees of symptoms, it was designed to detect change on an individual level. The authors advise that the score should not categorise individuals into groups (31,50,51) as this has no prediction on patient satisfaction post surgery. However, studies of the Oxford hip score have attempted to create a ranking system of excellent (>41), good (34-41), fair (27-33), or poor (<33) (52). Additionally when comparing pre and postoperative scores to look for changes in trends, all of the Oxford scores have been divided into equal patient number deciles for comparison (31,50).

In distinguishing between individuals with and without symptoms there are no cut off points. Presuming therefore that any one who was asymptomatic would have a perfect OSS has been proven to be incorrect. In a study Younis et al reported OSS scores ranging between 13 and 48 for asymptomatic individuals (41). This recording of positive scores in asymptomatic patients would potentially weaken any association found with pathology and symptoms.

The numeric rating scale (NRS)

The numeric rating scale is one of 3 commonly used single-dimensional pain-rating scales along with the visual analogue scale (VAS) and the verbal rating scale (VRS), which are all used to measure pain intensity. It is an 11-point scale representing a scale form no pain though to the worst possible pain, and contains 10 discrimination levels. Correlation between all three of these scores has been demonstrated (45). All have been demonstrated to be valid reliable and responsive.

The discriminate validity of an 11-point scale for the assessment of pain has been demonstrated by Jensen et al. Like many other scale scoring systems though, it is not recommended to try to categorise these into severity groups, rather the VRS should be adopted. However studies of cancer patients have compared the NRS to the VRS and concluded mild pain was quantified as being between 1 and 4, moderate between 5 and 6, and severe anything greater. Other studies however have shown that although well correlated there were wide ranges of VAS and NRS scores corresponding to each VRS category (44).

Application of the NRS to this study

1. Content validity

The NRS was asked for each shoulder and asked to describe the worst pain over the last 4 weeks. This fits biologically with our model that rotator cuff tears are a chronic condition, and that pain may fluctuate from time to time. However it asks only of maximal pain, and as we are reporting only the severest of symptoms and

not the average, we may draw a false increase in the association and severity of symptoms.

The content validity of the NRS is weak in comparison to the OSS as it only assesses pain, and of that, only one dimension of pain. It has been validated only to represent a patient's interpretation of their pain intensity. Unlike the OSS is does not distinguish between different types of pain and thus if only the worse are reported then again a false causality between symptoms and pathology may be enhanced.

In terms of the target population the NRS had been validated and is reliable for older generations including those with cognitive impairment. This would be a strength over the OSS.

2. Construct validity

The construct validity is good. The NRS demonstrates good convergent validity with other pain rating scales. It also is a patient based tool and not subject to clinician bias. Strength over the OSS is that it is easy to comprehend and has been reported to have a failure rate of less than 2%. Thus there is little scope for bias to occur.

In the Chingford study it is specifically asked with respect to the shoulder and thus should not be susceptible to bias from noise of surrounding areas. Although validated to be used for musculoskeletal pain, in particular postoperative pain, no studies have directly looked at its application to rotator cuff disease.

3. Discriminate validity

The discriminate validity is perhaps stronger for the NRS compared to the OSS. A 10 or 11-point scale has been validated to show it has enough sensitivity to allow reporting of changes in pain levels. It is also perhaps clearer to patients in determining those with and those without symptoms. Like the OSS it can be tempting to attempt to quantify the rank scores into groups. A number of studies have attempted to do this with the NRS and VAS compared to the VRS. Although good correlations were found wide ranges of scores were present within each VRS group demonstrating again that categorisation should be avoided.

Binary clinician asked question

This was asked by the examining clinician on the day of the ultrasound and referred specifically to that day; the answers were binary.

1. Content and discriminate validity

The content validity of using a binary question is relatively poor. As a single question this covers only one dimension of pain. As a binary answer it means that any discriminate validity to the severity of symptoms is lost. However it is a simple question to answer even in those with cognitive impairment, and is subject to little bias from surrounding areas.

Asking only about pain on the day does not well reflect the biology of rotator cuff tears, which are a chronic condition. If a patient is not experiencing pain that day

or perhaps only experiences night pain then any association between pathology and symptoms would be underestimated.

2. Construct validity

Bias may be introduced as this was a clinician led question rather than patient reported. Thus in the context of the interview it may be possible for the clinician to bias the answer one way or the other depending on other questions and examination findings. This bias would strengthen any association between pathology and symptoms. In order to reduce this bias the question would have to be always asked at the beginning before any PROMs were reviewed or any examination performed: This was not the case. However, on the other hand being clinician led may reduce bias from surrounding areas. For example a patient may perceive neck pain or thoracic pain to be shoulder pain.

Retrospective pain recollection

Patients were asked to recall if they had ever experienced shoulder pain and whether they had ever sought a medical consultation regarding this.

1. Content and discriminate validity

The content validity of using a binary question is relatively poor. As a single question this covers only one dimension of pain. As a binary answer it means that any discriminate validity to the severity of symptoms is lost.

Asking only about previous pain however, does potentially reflect the biology of rotator cuff tears, which are a chronic condition. Patients may have experienced pain in the past when a tear has taken place that has subsequently resolved.

2. Construct validity

Here lies the major potential weakness due to recall bias. Pain is rarely recalled accurately when reporting past musculoskeletal symptoms, and this is exaggerated as the time duration extends. Here the question merely asks whether pain has ever been present and does not even quantify if so when. Studies have reported that up to 3 months pain can be reliably reported. However a study at 6 years showed that 72% did not recall pain 6 years after reporting it (almost 100% in those with no current pain and 45% in those with current pain). Other factors that influenced the recollection were the duration of pain and certain psychological characteristics including education personality traits and age. It is likely therefore, that in this study we will underestimate the presence of previous pain. This will be more pronounced in those who have no current pain, and those who experienced pain a long time ago. This could potentially weaken any associations found with previous pain and rotator cuff tears.

2: Aims and objectives

This thesis aims to determine whether there is an association between rotator cuff tears, pain and functional loss in the general population. The hypothesis states that with increasing tear stage severity symptoms will become firstly more prevalent and secondly more intense.

The OSS is the only shoulder specific PROM that was obtained during the Chingford Y20 cohort study. Therefore the aim of this preliminary study is to validate its use as a primary outcome measure. Firstly, to validate splitting the score into pain and function sub scores, and secondly its use as a dichotomized score. Lastly to determine whether any of the OSS pain specific questions were more sensitive to capturing symptoms.

3: Patients and methods

SETTING AND SUBJECTS

The study participants were selected from the Chingford Study (as described in section 1.1). With consent and ethical approval 463 individuals who voluntarily attended the 20-year follow up visit completed a musculoskeletal questionnaire including the OSS, NRS, and asked whether they had shoulder pain at that moment, whether they had ever had shoulder pain, and whether they had ever sought medical advice for shoulder pain.

STATISTICAL ANALYSIS.

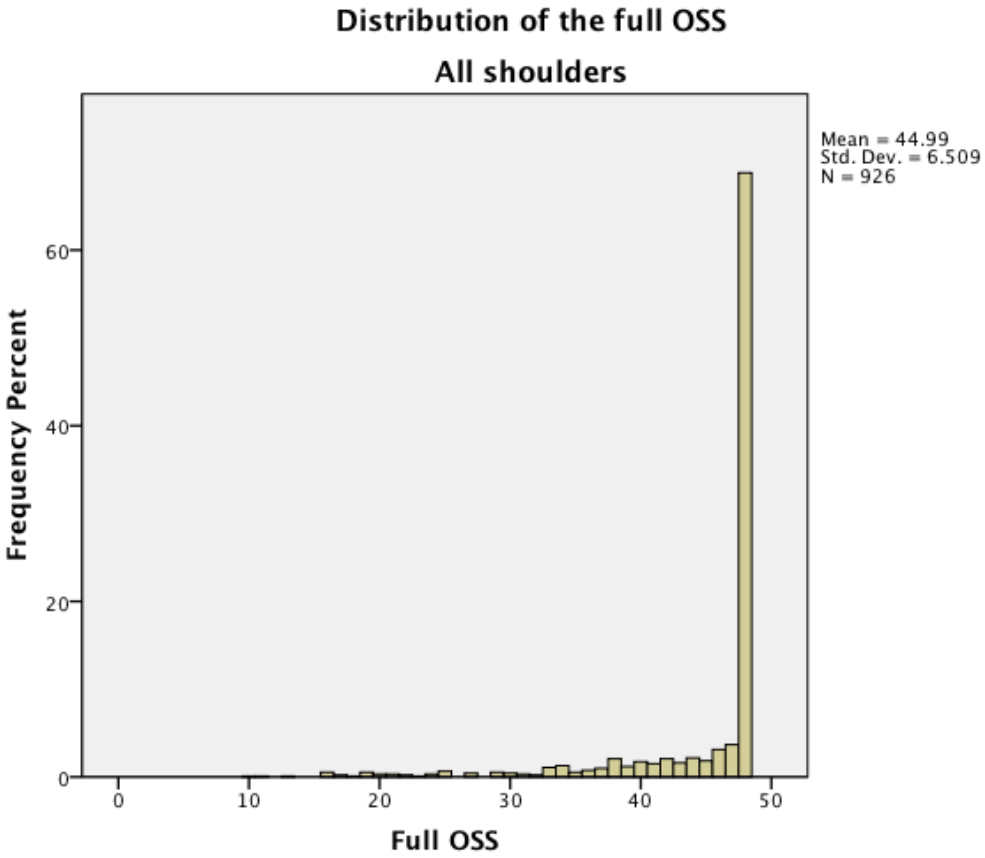
Scatterplots of the OSS, OSS pain sub score, OSS function sub score, and NRS were produced and correlations determined using Pearson correlation co-efficient. NRS, OSS and OSS pain sub scores were compared between those reporting pain on the day and those not. Pearson's correlation co-efficients were calculated. Box and whisker plots were produced between each of the specific pain questions and the NRS, and Pearson's correlation co-efficient determined.

All statistics were performed using IBM SPSS Statistics version 22 (IBM, Armonk, NY, USA).

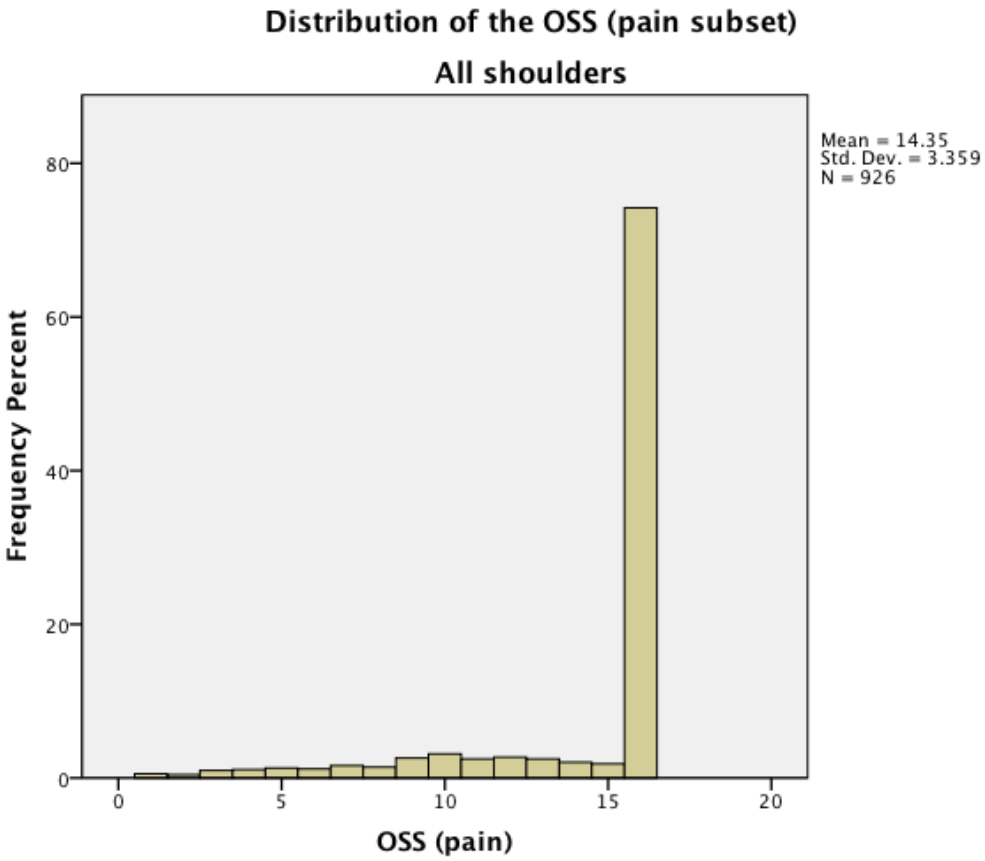
4: Results

1: Descriptive statistics

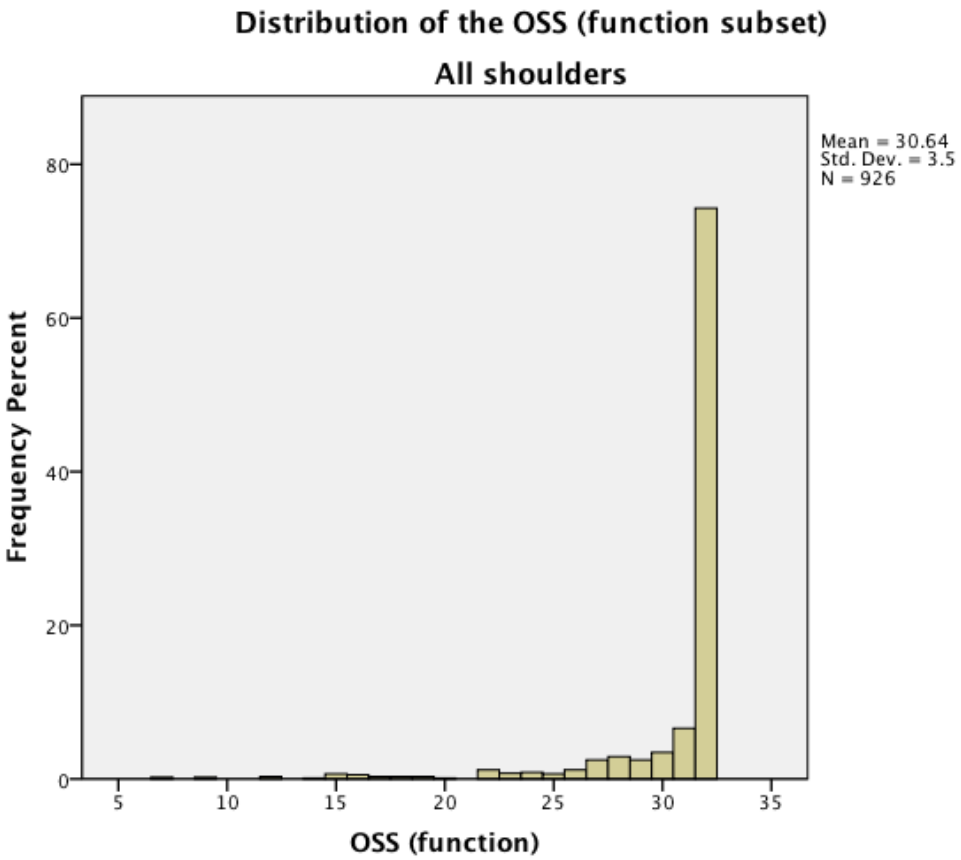
926 shoulders were included in the analysis. All OSS and NRS scores were abnormally distributed and negatively skewed due to the high number of asymptomatic shoulders. The median OSS was 48, OSS pain 16, OSS function 32, and NRS 1. The distributions are shown in the histograms 8-11.



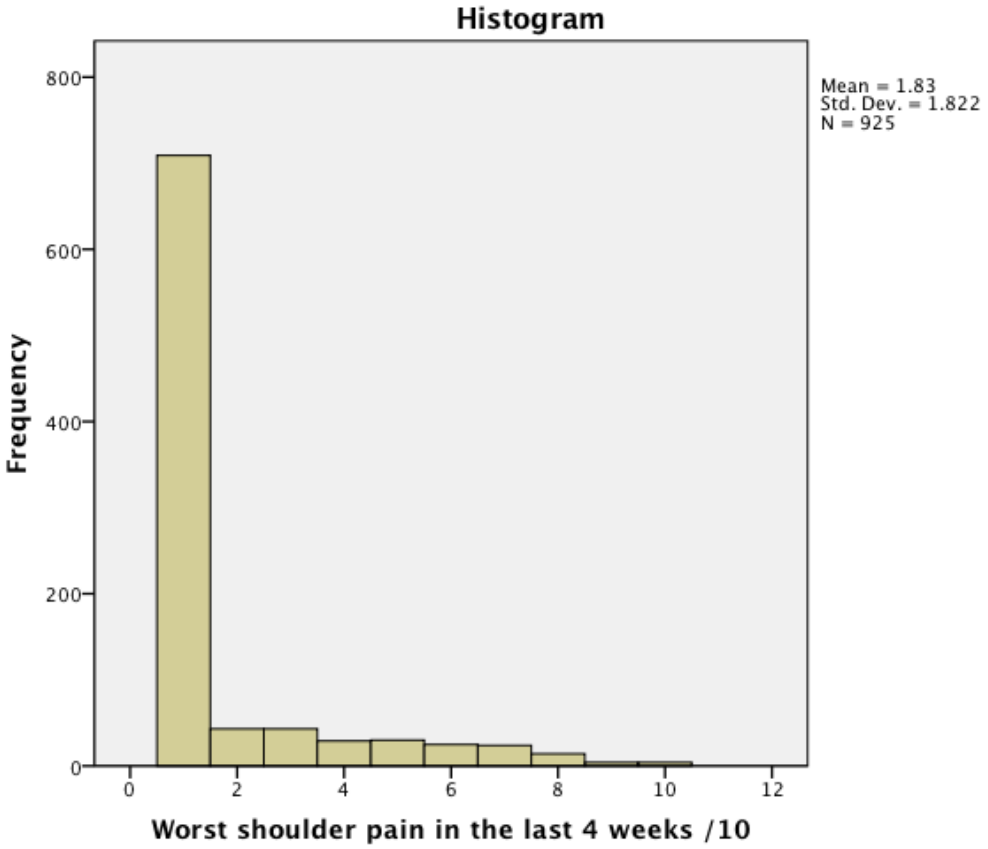
Graph 8: Abnormal distribution of full Oxford Shoulder Scores across the cohort



Graph 9: Abnormal distribution of the pain subset of the Oxford Shoulder Score



Graph 10: Abnormal distribution of the function subset of the Oxford Shoulder Score



Graph 11: Abnormal distribution of the NRS scores across the cohort

187 (20.2%) of shoulders were painful on the day of examination, and of these 97 (51.9%) had sought medical advice regarding this pain. 341 (36.9%) shoulders had had previous pain, and of these 180 (52.8%) had sought medical advice for the pain. Overall 180 (19.4%) of all shoulders have had pain resulting in the seeking of medical advice. This is shown in table 7.

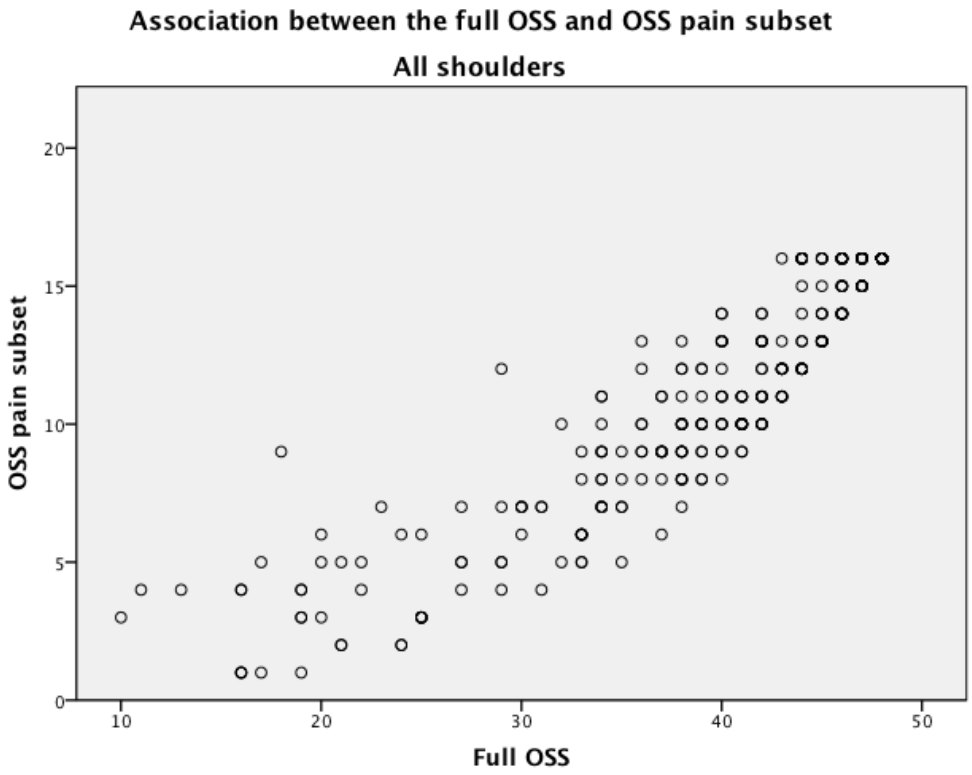
Number of Individuals with shoulder pain, past or present, and the proportion seeking medical advice for this

		Current pain		Previous pain		GP	
		Yes	No	Yes	No	Yes	No
Current pain	Yes	187		187	0	97	90
	No		739	154	583	83	651
Previous pain	Yes	187	154	341		180	158
	No	0	583		583	0	583

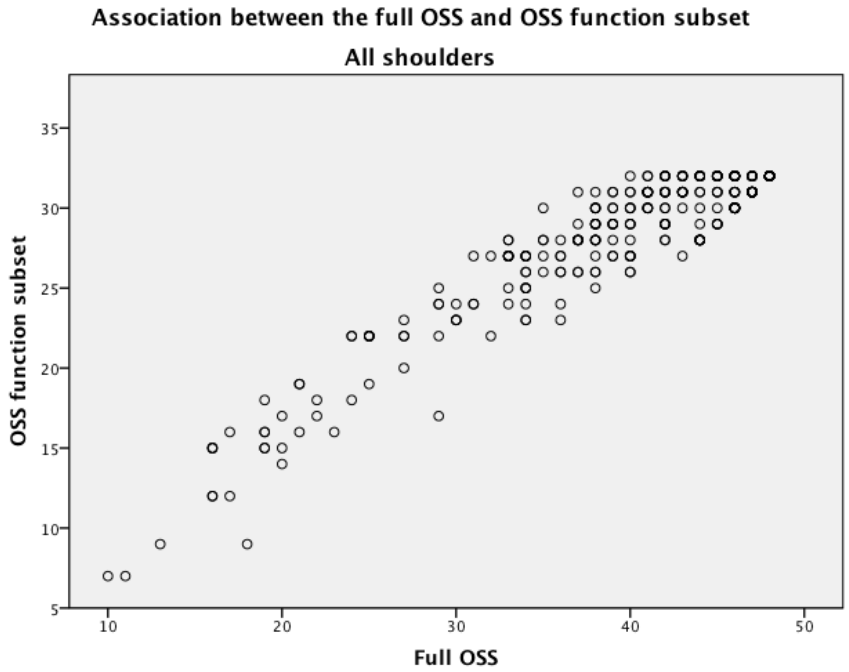
Table 7: Number of symptomatic shoulders (N=926)

2: Is it valid to Split the OSS into two subscales?

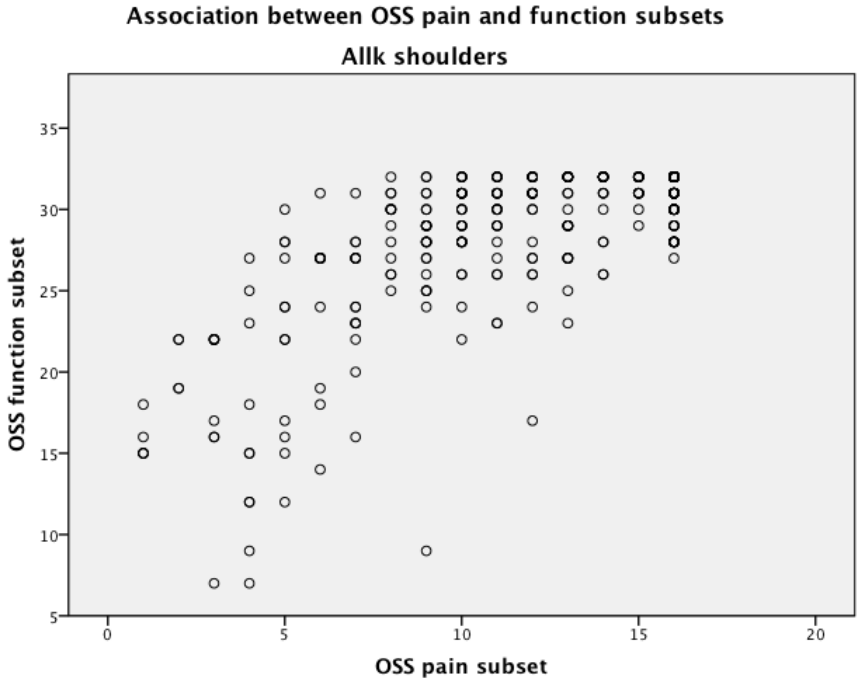
Pearson's correlation between the OSS and OSS pain was 0.947 (p<0.001, 95%CI 0.940-0.953), between the OSS and OSS function was 0.951 (p<0.001, 95%CI 0.944-0.957), between the OSS pain and OSS function was 0.801 (p<0.001, 95%CI 0.777-0.823). See graphs 12-14.



Graph 12: Correlation between the OSS and OSS pain subset. R=0.947, p<0.001

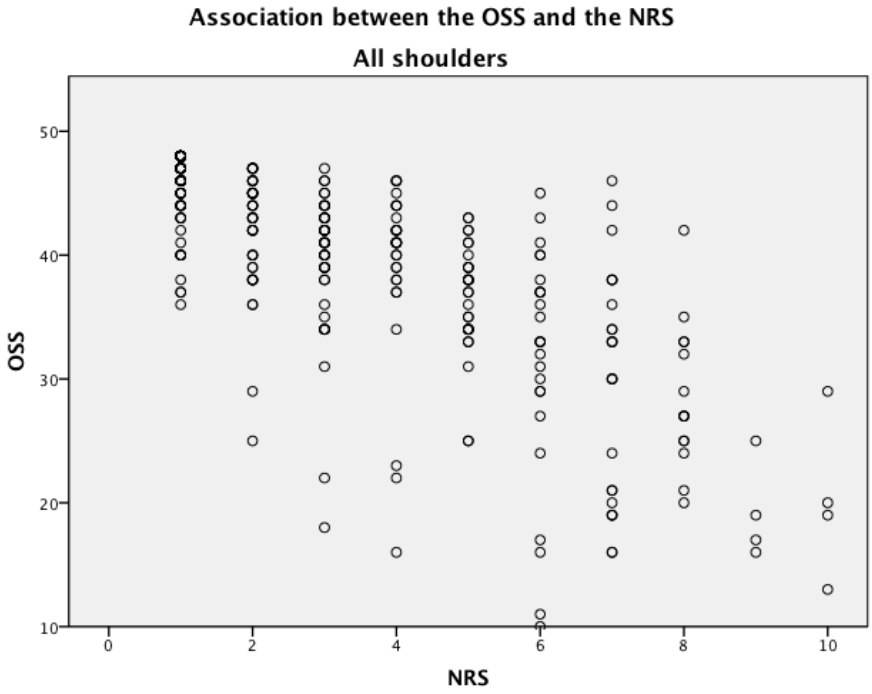


Graph 13: Correlation between the OSS and OSS function subset. R=0.951, p<0.001

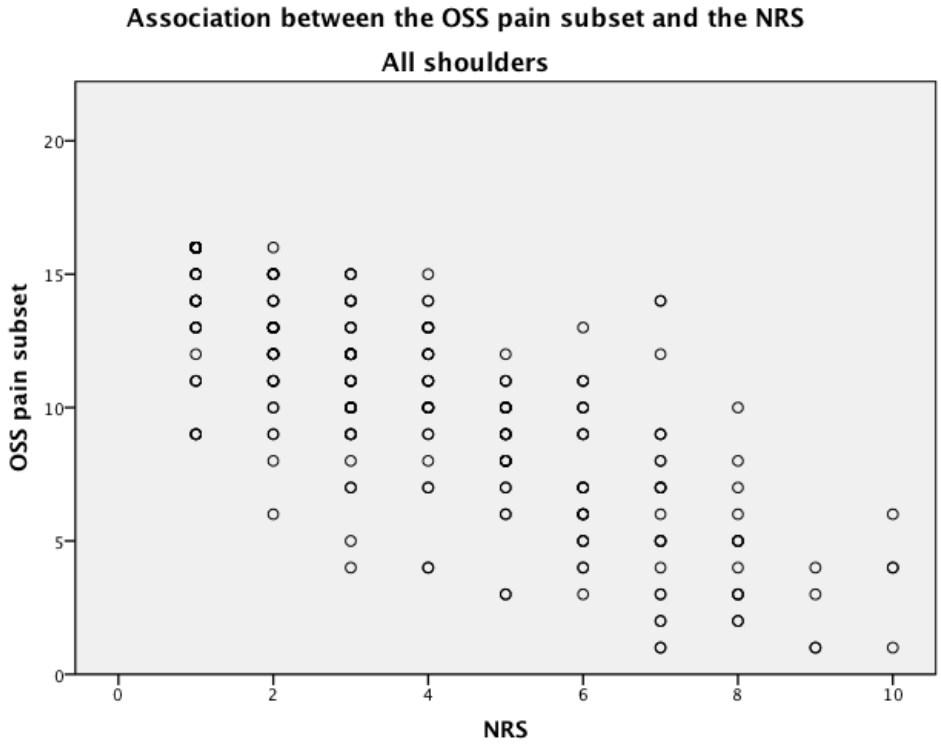


Graph 14: Correlation between the OSS pain and function subsets. R=0.801, p<0.001

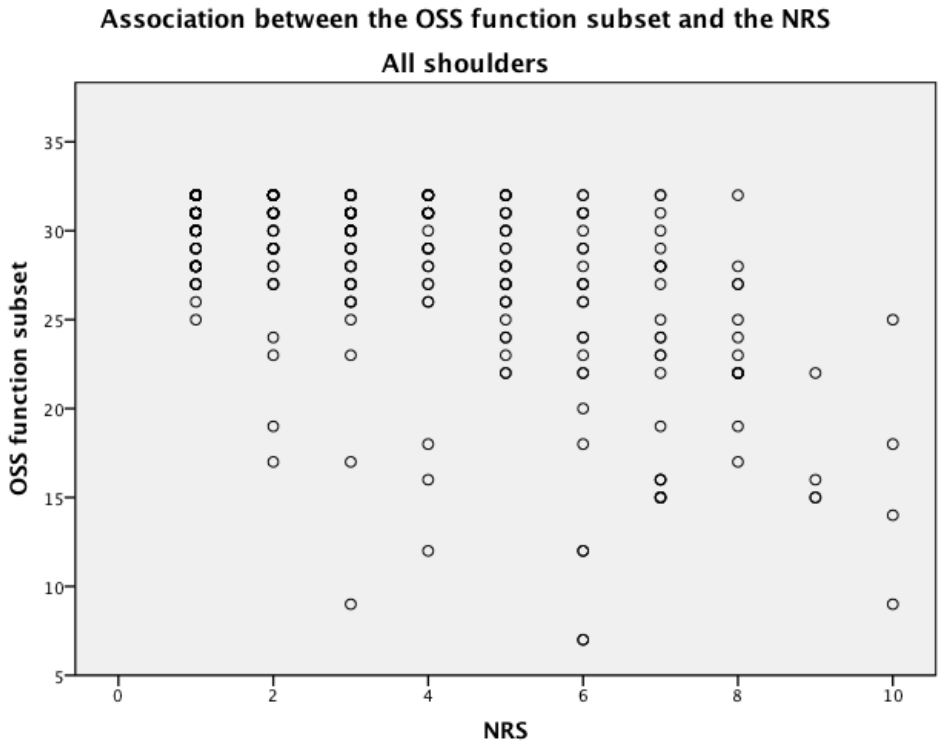
Pearson's correlation between the NRS and OSS was 0.851 (p<0.001, 95%CI 0.832-0.868), between the NRS and OSS pain was 0.901 (p<0.001, 95%CI 0.888-0.912), between the NRS and OSS function was 0.719 (p<0.001, 95%CI 0.686-0.749), See graphs 15-17. When OSS pain scores were divided into groups with perfect scores being considered asymptomatic and the remaining scores divided into 3 equal tertiles perfect correlation was found with the NRS in 88.3% of cases. Only 3.7% of the whole cohort disagreed when only comparing perfect scores. See table 8.



Graph 15: Correlation between the OSS NRS. R=0.851, p<0.001



Graph 16: Correlation between the OSS pain subset and the NRS. R=0.901, p<0.001



Graph 17: Correlation between the OSS function subset and the NRS. R=0.719, p<0.001

Correlation between the OSS pain subset and the NRS when scores were grouped relative to severity

OSS (pain subset) binned into 4 groups		NRS binned into 4 groups				Total
		No symptoms (1)	Mild symptoms (2-3)	Moderate symptoms (4-5)	Severe symptoms (6-10)	
No symptoms (16)	Count	685	1	0	0	686
Mild symptoms (12-15)	Count	17	50	13	4	84
Moderate symptoms (9-11)	Count	7	28	28	13	76
Severe symptoms (0-10)	Count	0	7	18	54	79

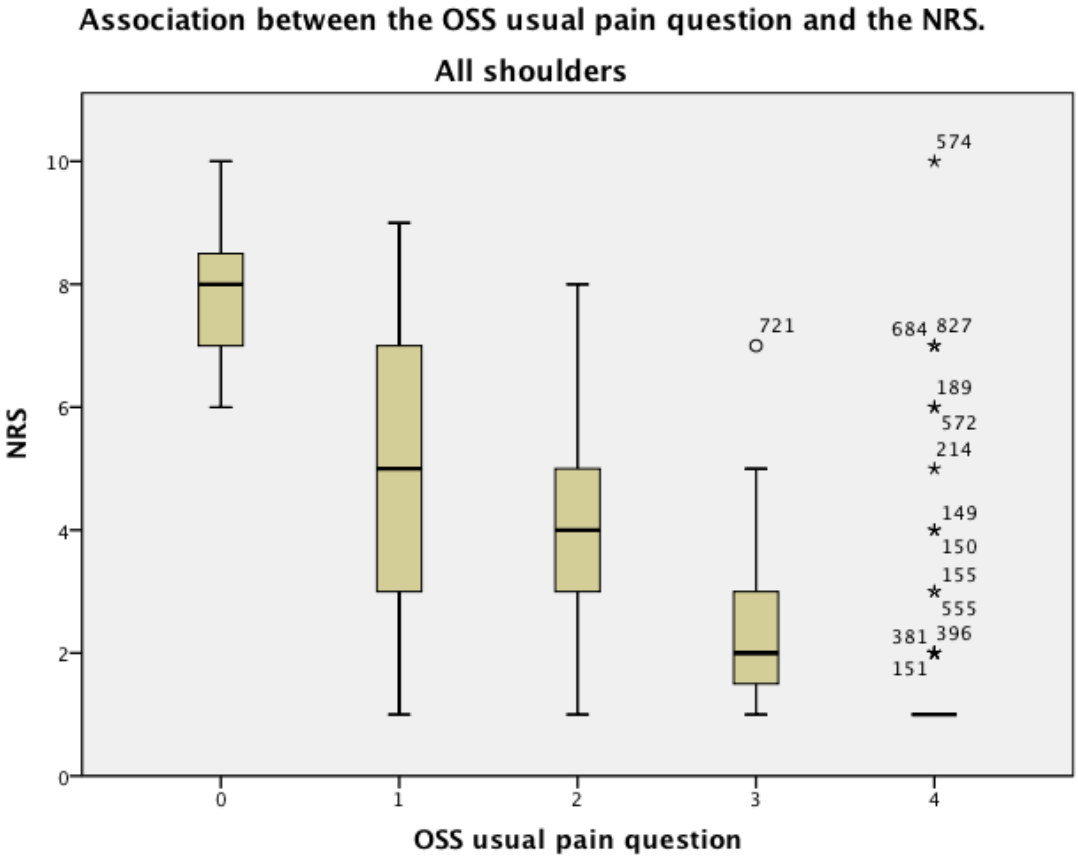
Table 8: Agreement in scores was found in 88.3% of cases

ASSOCIATION BETWEEN BINARY AND CONTINUOUS PAIN ASSESSMENTS

Pearson's correlation between the NRS and binary pain question was 0.816 ($p<0.001$, 95%CI 0.793-0.836), between the OSS pain and binary pain question was 0.812 ($p<0.001$, 95%CI 0.789-0.833), between the OSS and binary pain question was 0.759 ($p<0.001$, 95%CI 0.730-0.785).

3: ANALYSIS OF SEPARATE PAIN QUESTIONS IN THE OSS:

Pearson's correlation between the NRS and night pain question was 0.761 ($p<0.001$, 95%CI 0.732-0.787), between the NRS and worst pain question was 0.787 ($p<0.001$, 95%CI 0.761-0.810), between the NRS and work pain question was 0.689 ($p<0.001$, 95%CI 0.654-0.721). Strongest correlation was found between the NRS and usual pain question which was 0.803 ($p<0.001$, 95%CI 0.779-0.825), See graph 18. Overall there was no question that had significantly better correlation with the NRS.



Graph 18: Correlation between the NRS and the OSS usual pain question was R=0.803 (p<0.001), though no specific question gave better correlations.

5: Discussion

In order to answer whether rotator cuff tears are associated with pain and functional loss the OSS (full) will be dichotomized and used as the primary outcome measure. Secondary measures of the dichotomized OSS (pain) and dichotomized OSS (function) will be used to determine differences between pain and functional losses.

Although the question requires a binary definition of pain and functional loss, and we have a binary pain question, we feel that the biology of rotator cuff tears

is chronic and therefore is not well represented merely by asking of pain on a specific day. The NRS and the OSS offer better content validity by reflecting the biology of rotator cuff tears in asking after symptoms in the last four weeks. The OSS offers still increased content validity over the NRS by asking of different aspects of pain. By nature of the OSS being a PROM we feel that this offers the greatest construct validity of the scales used. The weakness of the OSS however resides in categorising individuals by score into symptomatic and asymptomatic groups. This study has shown though that both the NRS and the OSS (pain) showed good correlation with the binary pain question, notwithstanding that in fact the questions were being asked over different time frames which would add bias against correlation. The other limitation of the OSS is that although designed to be joint specific with the use of hypothetical questions, these can be misinterpreted, and this would become more so in an ageing cohort. However good correlation was shown in this study between the OSS pain score and the NRS even when split into groups.

To determine how to define symptomatic and asymptomatic, we defined having symptoms as any non-perfect OSS. This was on the rationale that the individuals that because of their shoulder they were unable to do an activity to the full, or had some pain at a given time. This is supported by the correlations between the OSS pain subset and the NRS when symptoms were lumped into groups. Perfect correlation was achieved in 96.7% of cases when only those with either perfect OSS pain scores or perfect NRS scores were compared. Within all groups the correlation between the 2 was perfect in 88.3% of cases. There have however been studies in the literature that have attempted to look at the minimal

clinically important changes in scores. With respect to the Oxford shoulder score Van Kampen et al (53) found that minimal change in satisfaction required a 6-point change in the OSS. Similarly Clement et al. found a 5-point change using the Oxford Knee Score (54). Pain scales too have been evaluated, and Farrar et al. determined that on a 1-10 pain scale a 2 point change was considered significant to the individual, or in other scales a 33% change(55,56). However, all of these are looking for what is deemed important in changes of score is response to intervention rather than what defines symptoms. This study requires a definition of symptoms. We therefore have decided to use a single point of change as our main primary outcome as this indicates any symptoms from mild to incapacitating. In keeping with other studies we also decided to run a second analysis where we will use a three-point change in the OSS, or OSS subsets to attempt to capture the groups where symptoms may be of clinical relevance. 3 points was chosen after discussion with one of the creators of the OSS, and also to reflect other similar works. For example Yamaguchi et al defined their symptoms using a 3-point change on the VAS. Furthermore this is in keeping with studies which have tried to determine what is a patient acceptable state (PAS) in pain scores such as the NRS and VAS. A study by Tashjian et al. determined this to be a 3-point change (57). Although PASs have not been determined and are not recommended to use in scores such as the OSS due to the wide range of scores, we have shown that there is strong correlation between the NRS and OSS pain subset, which would add validity to our use in this way.

To determine differences in symptoms in terms of pain and function, the secondary outcomes used will be the dichotomized OSS pain and function

subscales. Although the OSS has not formally been validated to be split, this is justified by the stronger correlations of the OSS (pain) to the NRS than the OSS (function). Correlation coefficients between the OSS pain subset and the NRS taken from my dataset were 0.901 (95%CI = 0.888-0.912), and between the OSS function subset and the NRS 0.699 (95%CI = 0.686-0.749). Correlation between the full OSS and the NRS was 0.851 (95%CI = 0.832-0.868). To justify its use in its dichotomized form good correlations were shown compared to the primary pain question, which were good equal to the NRS.

The use of individual questions of the OSS pain subset is not justified from our data. No particular question has significantly greater correlation to the NRS in our cohort, and each question correlated well.

A further analysis will be performed based upon health economics. Outcomes would include firstly binary retrospective pain recollection and secondly previous GP attendances. Although retrospective data collection is subject to considerable bias we feel that it reflects the biology of the condition well and thus gives useful information. A health economic slant is added by determining whether the GP has been visited due to shoulder pain and highlights the importance of the symptoms to the individual. Data interpretation however will need to account for the potential recall bias of such questions and appreciate that results may be underestimated.

In order to determine whether the degree of pathology correlates with the degree of symptoms we will use the full OSS as the primary outcome measure, and the OSS pain and function subscales as secondary measures. The OSS allows the assessment of pain and function in one single score, as well as analysing them independently as justified above. Using the OSS in its entirety as a continuous variable will however cause difficulties with statistical analysis due to the significant negative skew of the data. Thus when analysis of symptom severity takes place all normal scores will be removed and results will only draw comparisons between different tendons groups in shoulders that are symptomatic. The rationale for this and statistical transformations used are outlined in the first appendix, section 3.1.

6: Summary

This study has provided justifications for our choice of the use of the OSS in this study in terms of content and construct validity.

The outcomes used in this study to define symptoms will therefore be:

1. To determine symptom likelihood
    a. Primary outcome measure:
        i. Dichotomised OSS with a change of 1-point defining symptoms
        ii. Dichotomised OSS with a change of 3-points to reflect PAS
    b. Secondary outcome measures:
        i. Dichotomised OSS pain and function subscales with a change of 1-point defining symptoms

ii. Dichotomised OSS pain and function subscales with a change of 3-points to reflect PAS

2. To determine symptom severity in all symptomatic shoulders only
    a. Primary outcome measure:
        i. OSS (all perfect scores removed)
        ii. OSS pain and function subscales (all perfect scores removed)

Section 2: Study Results

1: Prevalence of rotator cuff tears in the general population

1: Background information

Rotator cuff tears are the most common shoulder disorder treated by orthopaedic surgeons. The shoulder itself is the third most common site of musculoskeletal pain (1), with an estimated 20% of the population reporting pain at any given time (2). Due to the high disability rates and thus costs (3,4) there is a wealth of literature on the management of rotator cuff tears, and there are many different treatments options, none of which are standardised. One of the possible reasons for this lack of management consensus is the significant variation in clinical manifestations of rotator cuff tears and the lack of understanding of the basic epidemiology and natural history of rotator cuff tears.

The published literature lacks consensus on the population prevalence of full-thickness rotator cuff tears. Wide variations in prevalence have been estimated from cadaveric and radiological studies in both symptomatic and asymptomatic shoulders. Reilly et al. (58) performed a systematic review of the prevalence of rotator cuff tears according to cadaveric and radiological studies.

Of the 30 cadaveric studies included in the systematic review, the overall prevalence was 12.7%, ranging from 2-40%. Cadaveric studies, however, are not a good representation of the general population. Although they provide gold standard anatomy and will include both tears that were asymptomatic and

asymptomatic they also are likely to contain a significantly older cohort of shoulders, and thus as rotator cuff tears are associated with age (59-62), cigarette smoking (63), and degenerate histo-pathological changes in tendons (64), such studies would likely overestimate population prevalence. Furthermore considerable variance in estimated prevalences will come from the different cohorts studied, which are often not specified.

11 ultrasound studies were included, of which 2 included asymptomatic shoulders, and the remainder studied symptomatic shoulders. The average prevalence in the asymptomatic group was 38.9%, and 41.4% for the symptomatic shoulders. 12 MRI studies, including 4 asymptomatic studies, and 8 symptomatic studies were included. The average prevalence for asymptomatic shoulders was 26.2% and 49.4% for symptomatic shoulders.

None of the studies included were population based. Furthermore, data from such ultrasound and MRI studies will not reflect true population prevalence as they vary due to the study type and specific cohort inclusion. Despite attempting to reveal the population prevalence by setting up studies using asymptomatic volunteers, selection bias is introduced by excluding symptomatic shoulders, and further recruitment bias is introduced in the cohort selection. Many of the asymptomatic studies have been recruited from orthopaedic outpatient clinics (60-62,65-67), and thus individuals have been pre-selected with musculoskeletal pathology. Other smaller cohorts have been derived from younger volunteers in the armed forces (68) or have been advertised for (69,70), and thus are self selecting, and do not represent the general population.

Although rotator cuff tears are not always symptomatic (60,61,71), if it is as our hypothesis suggests, that rotator cuff tears are associated with symptoms, studying only symptomatic shoulders would create a selection bias leading to an over reporting of prevalence. Likewise selecting only asymptomatic shoulders would have a bias towards under-reporting prevalence. Recognising this, studies have looked at cohorts of patients who are being treated for painful rotator cuff tears, and evaluated the contralateral pain free shoulder. These have determined a 35-50% increase in risk of asymptomatic rotator cuff tear (71). Although this provides evidence that not all tears are symptomatic and that even within an individual the clinical manifestation of a tear may differ, population prevalence cannot be derived, as there is association with the same genetic, and systemic risk factors that potentially would introduce detection bias.

Retrospective cohort studies have been performed on a small mountain population in Japan (72,73). Subjects volunteered for a medical check up for the early detection of cancer, and a cohort of 683 from these was randomly selected. The cohort has not been demonstrated to reflect the general population. Of this cohort 20.7-22.1% of shoulders were found to have full-thickness rotator cuff tears, but no comments were made on the population prevalence. No general population cohorts representative of the UK or the developed Western society have been investigated.

The prevalence of full-thickness tears is not the complete story. To achieve the ultimate goal of understanding the natural history of rotator cuff tears, it is also

important to know the epidemiology of tendon degeneration that potentially precedes a full-thickness tear. Estimations of partial-thickness tear prevalence range from 8.5%-38.9% depending on the modality of study used (58). The same selection and detection bias discussed with full-thickness tears holds true for studies detecting partial-thickness tears. Furthermore, although accurate diagnoses may be achievable from cadaveric studies, those relying on ultrasound imaging may be less accurate. High-definition ultrasound has been shown to achieve high levels of accuracy when diagnosing full-thickness tears but lower levels when diagnosing partial tears due to poorer sensitivities of around 72% (8-13,74), which constitutes a major limitation to such studies. High-definition ultrasound has not only made ultrasound a more cost-effective and accessible diagnostic tool but has also enabled the detection of potential pre-tear degenerate changes at the tendon enthesis, which can validly and accurately be detected (74) (See appendix 2 for full details). No studies have looked at the prevalence of abnormal tendons in any cohort of individuals.

There are no studies to date that have aimed to determine the prevalence of rotator cuff tendinopathy, from the abnormal enthesis up to full-thickness tears using a general population cohort. This epidemiology is an important foundation required prior to any understanding of the natural progression of rotator cuff tendinopathy and its clinical manifestations that will ultimately contribute to consistent management plans.

2: Aims and objectives

The aim of this chapter is to define the prevalence of different stages of rotator cuff tendinopathy in the general population.

3: Patients and methods

PATIENTS

The study participants were selected from the Chingford Study, a well described prospective population-based longitudinal study of osteoporosis, comprising initially of 1003 women, derived from the register of a large general practice in Chingford, North London (see section 1). The women aged 44-67 years at baseline are representative of women in the UK general population with respect to weight, height, and smoking characteristics. At year 20 of the study 464 underwent a shoulder musculoskeletal assessment including the Oxford shoulder score, and bilateral shoulder ultrasound (Of the original 1003, 158 women had died, 111 were unable to attend, 218 had either moved away, dropped out or were lost to follow up, and 52 attended the year 20 visit but did not have a musculoskeletal assessment due to absence of the assessor). The local ethics committee approved the study and consent was obtained from each woman (Outer North East London Research Ethics Committee (formerly Barking and Havering and Waltham Forest RECs), LREC (R&WF) 96).

METHODS

For each participant age, height measured in cm (to the nearest 0.1cm) in a standing position with shoes removed, using a wall-mounted stadiometer (Leicester height measure, Seca) and weight in kg (to the nearest 0.1kg) by electronic scales with shoes removed were recorded. A fixed SOPP (standard operating procedure protocol) was followed for each musculoskeletal assessment. This was followed by ultrasound assessment of the shoulders (left then right).

The ultrasound examination of the 464 women was performed using the GE voluson i portable ultrasound machine with a 10-16MHz linear probe and is detailed in section 2. 343 individuals were scanned by Ms Hinsley, Orthopaedic SpR and Author, and 121 by Mr Nichols, Orthopaedic SpR, both of whom underwent recognised training and validation studies prior to the study as recommended by the BESS focus group (for validity see section 1.2.4). A fixed protocol was derived according to the recommendations of the Nuffield Orthopaedic Centre Musculoskeletal radiology department (see section 1.2.3), and results were recorded on a predetermined data entry sheet. Tendons were classified into one of 4 working groups based upon ultrasound measurements:

1. Normal tendons
2. Abnormal tendons and partial thickness tears
3. Single tendon full-thickness tears (0-2.5cm)
4. Multi-tendon full-thickness tears (>2.5cm)

Justification of the development and use of the above tendon classification is detailed in section 1.2.

STATISTICAL METHODS

Age, BMI, and arm dominance, were compared across the four different tendinopathy groups. Wilcoxon rank sum test, one-way ANOVA, and chi-squared tests were used for non-normal, normal and categorical data respectively. Population prevalence of full-thickness tears was defined as having at least one unilateral full-thickness tear (or any kind of tendinopathy for this prevalence). This was calculated by summing the percentage with unilateral tears and the percentage with bilateral tears for each age group.

All statistics were performed using IBM SPSS Statistics version 22 (IBM, Armonk, NY, USA).

4: Results

1. DESCRIPTIVE STATISTICS

464 individuals (928 shoulders) were included in the study. Age was not normally distributed and the median age was 71 (range 65-84). BMI was normally distributed with a mean of 27.8. Of the 516 individuals attending the year 20 visit, the median age was 71 (range 63-85), and mean BMI was 27.8. There was no statistical difference in age or BMI between those who underwent a shoulder assessment and the complete cohort attending the year 20 visit ($p=0.222/p=0.136$).

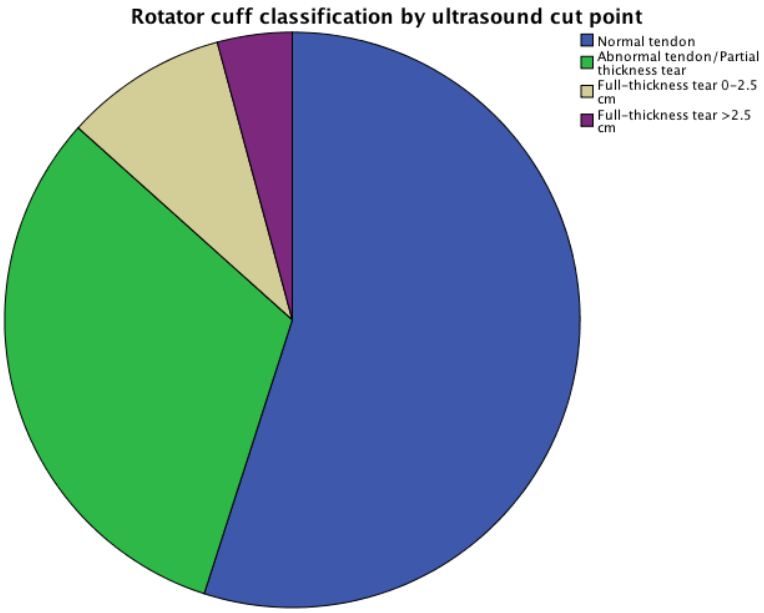
Of the 928 shoulders 510 were diagnosed as normal on ultrasound, 294 abnormal or partial-thickness tears, and 85 full-thickness tears (0-2.5cm), and

39 full-thickness tears (>2.5cm) (see table 9 and graph 19). The distribution of age across each tendinopathy group was significantly different (Kruskal Wallis Test p<0.001). There was no statistical difference in BMI for each tendinopathy group (One way ANOVA p=0.080). 840 (90.5%) were right-handed, and 88 (9.5%) were left-handed. There was no difference in right and left-handed individuals in each tendinopathy group (Chi^2 test p=0.771)

Demographics of shoulders included in the study

	Frequency	%	Median age	Mean BMI	Dominant arm (%)
Normal	510	55.0%	70	27.5	46.1%
Abnormal/ Partial tear	294	31.7%	73	28.0	52.7%
Full-thickness tears 0-2.5cm	85	9.2%	74	27.9	58.8%
Full-thickness tears >2.5cm	39	4.2%	74	29.6	61.5%
All	928		71	27.8	50%

Table 9: Increasing age and dominant arms were associated with increasing pathology



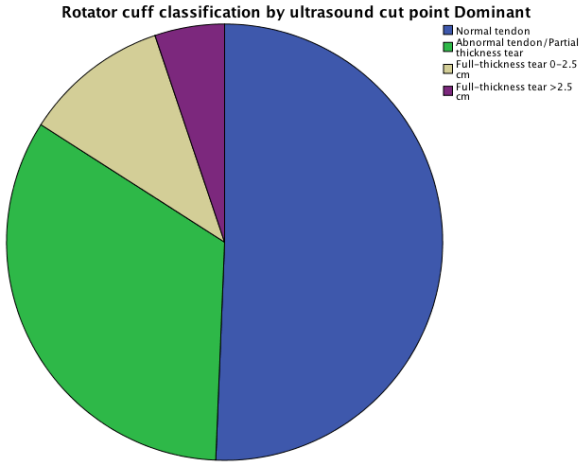
Graph 19: Rotator cuff tendinopathy prevalence

In the dominant arm 235 were diagnosed as normal on ultrasound, 155 abnormal or partial-thickness tears, 50 full-thickness tears (0-2.5cm), and 24 full-thickness tears (>2.5cm). In the non-dominant arm 275 were diagnosed as normal on ultrasound, 139 abnormal or partial-thickness tears, 35 full-thickness tears (0-2.5cm), and 15 full-thickness tears (>2.5cm) (see table 10 and graphs 20-21). There was a statistical difference in proportions of dominant and non-dominant arms in each tendinopathy group (Chi^2 test p=0.033), with there being significantly more non-dominant arms in the normal tendon group (Chi^2 test p=0.010), and significantly more dominant arms in those with full-thickness tears (Chi^2 test p=0.026).

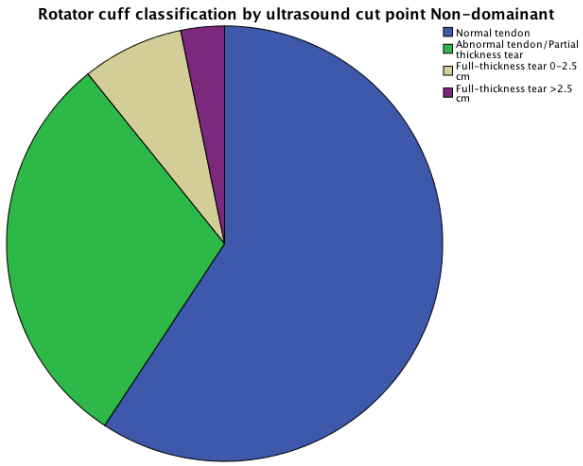
Demographics of shoulders included in the study

	Dominant Arm		Non-dominant arm	
	Freq.	%	Freq.	%
Normal tendon	235	50.6	275	59.3
Abnormal tendon / Partial-thickness tear	155	33.4	139	30.0
Full-thickness tear (<2cm)	50	10.8	35	7.5
Full-thickness tear (>2cm)	24	5.2	15	3.2
Total	464	100.0	464	100.0

Table 10: frequency of tendinopathy groups in the dominant and non-dominant arms



Graph 20: Rotator cuff tendinopathy, Dominant arm: Showing a smaller number of normal tendons compared to the non-dominant arm



Graph 21: Rotator cuff tendinopathy, Non-dominant arm, showing greater number of normal tendons compared to the dominant arm

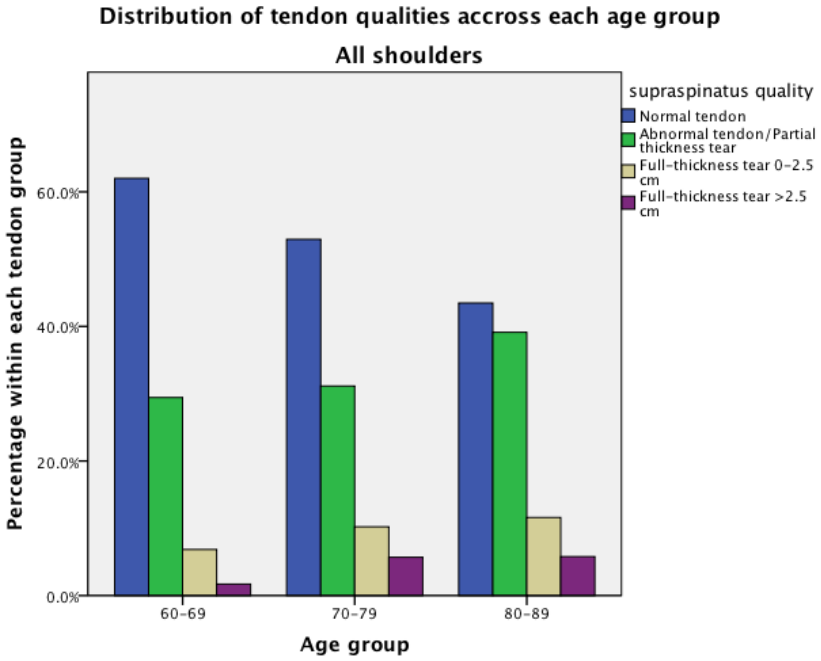
2: PREVALENCE

The prevalence of rotator cuff tendinopathy in the dominant and non-dominant arms according to each age decile is shown in table 11 and graphs 4-6. The distribution of tendinopathy differed between age groups (Dominant arm Kruskal-Wallace p=0.002; non-dominant arm p=0.037) with more pathology in the older age groups. There was more pathology in the dominant compared to non-dominant arms (Mann-Witney-U p=0.004)(Or if assume paired, Wilcoxon rank test p<0.001). There was no difference in tendinopathy distributions in BMI groups or right and left-handers.

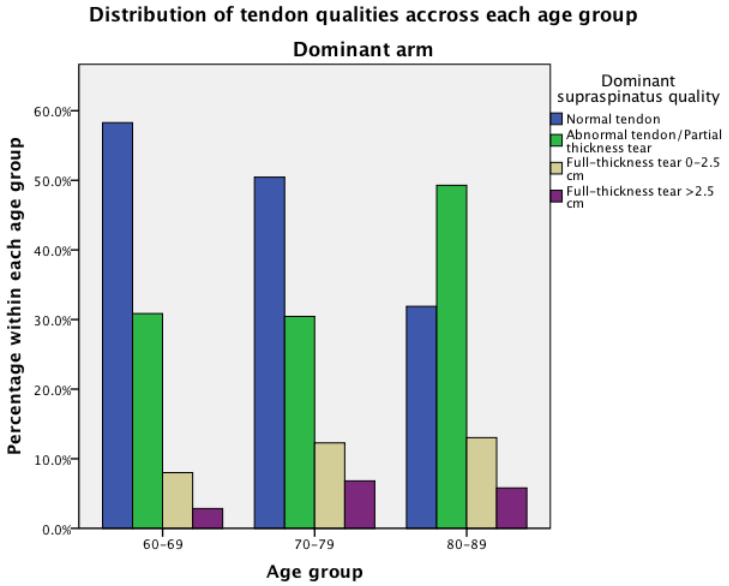
Prevalence of rotator cuff tendinopathy according to age decile and arm dominance

Dominant arm	Age groups at time of shoulder exam							
	60-69		70-79		80-89		Total	
	Count	%	Count	%	Count	%	Count	%
Normal tendon	102	58.3%	111	50.5%	22	31.9%	235	50.6%
Abnormal tendon/Partial thickness tear	54	30.9%	67	30.5%	34	49.3%	155	33.4%
Full-thickness tear 0-2.5 cm	14	8.0%	27	12.3%	9	13.0%	50	10.8%
Full-thickness tear >2.5 cm	5	2.9%	15	6.8%	4	5.8%	24	5.2%
Total	175	100.0%	220	100.0%	69	100.0%	464	100.0%
Non-dominant arm	**Age groups at time of shoulder exam**							
	60-69		70-79		80-89		Total	
	Count	%	Count	%	Count	%	Count	%
Normal tendon	115	65.7%	122	55.5%	38	55.1%	275	59.3%
Abnormal tendon/Partial thickness tear	49	28.0%	70	31.8%	20	29.0%	139	30.0%
Full-thickness tear 0-2.5 cm	10	5.7%	18	8.2%	7	10.1%	35	7.5%
Full-thickness tear >2.5 cm	1	.6%	10	4.5%	4	5.8%	15	3.2%
Total	175	100.0%	220	100.0%	69	100.0%	464	100.0%

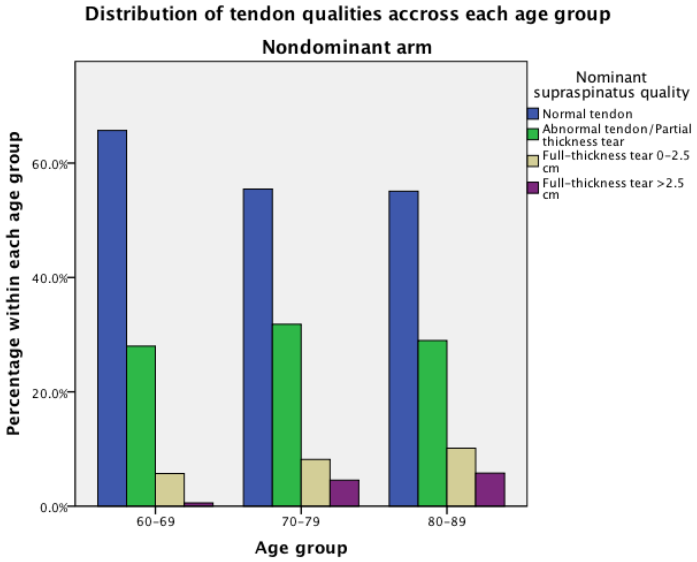
Table 11: Normal tendons are more prevalent in the younger age groups and the non-dominant arm



Graph 22: Normal tendons become less prevalent with increasing age



Graph 23: Normal tendons are less prevalent with increasing age and less prevalent in the dominant arm than the non-dominant arm



Graph 24: Normal tendons are less prevalent with increasing age and less prevalent in the dominant arm than the non-dominant arm

Uni-variable analysis demonstrated the increased relative risk of full thickness tear was 1.48 (95%CI 1.06-2.07, p=0.026) in the dominant compared to non-dominant arm; 1.86 (95%CI 1.24-2.78, p=0.002) and 2.03 (95%CI 1.23-3.34, p=0.010) for age groups 70-79 and 80-89 respectively compared to 60-69.

Multi-variable binary logistic regression showed that both increasing age group and being the dominant arm increased the likelihood of full-thickness tear (see table 12).

Relative risks of full-thickness tear

	RR	OR	95% CI	significance
Constant		0.073	-	<0.001
Age group		-	-	0.004
Age 70-79	2.072	2.026	1.286-3.190	0.002
Age 80-89	2.293	2.256	1.264-4.027	0.006
Dominant arm	1.640	1.580	1.073-2.326	0.021

Table 12: Relative risks of full-thickness tear increased with age group and in the dominant arm when compared to those aged 60-69 with non-dominant arms

For all tendon abnormalities uni-variable analysis demonstrated the increased relative risk of tendon abnormality was 1.21 (95%CI 1.05-1.40, p=0.010) in the dominant compared to non-dominant arm; 1.24 (95%CI 1.04-1.46, p=0.011) and 1.49 (95%CI 1.22-1.81, p<0.001) for age groups 70-79 and 80-89 respectively compared to 60-69.

Multivariable binary logistic regression showed that both increasing age group and being the dominant arm increased the likelihood of tendon abnormality (see table 13).

Relative risk of tendon abnormalities

	RR	OR	95% CI	significance
Constant		0.511	-	<0.001
Age group		-	-	0.001
Age 70-79	1.850	1.454	1.091-1.937	0.011
Age 80-89	2.390	2.133	1.428-3.187	<0.001
Dominant arm	1.826	1.426	1.098-1.853	0.008

Table 13: Relative risks of tendon abnormality increased with age and in the dominant arm when compared to those aged 60-69 with non-dominant arms

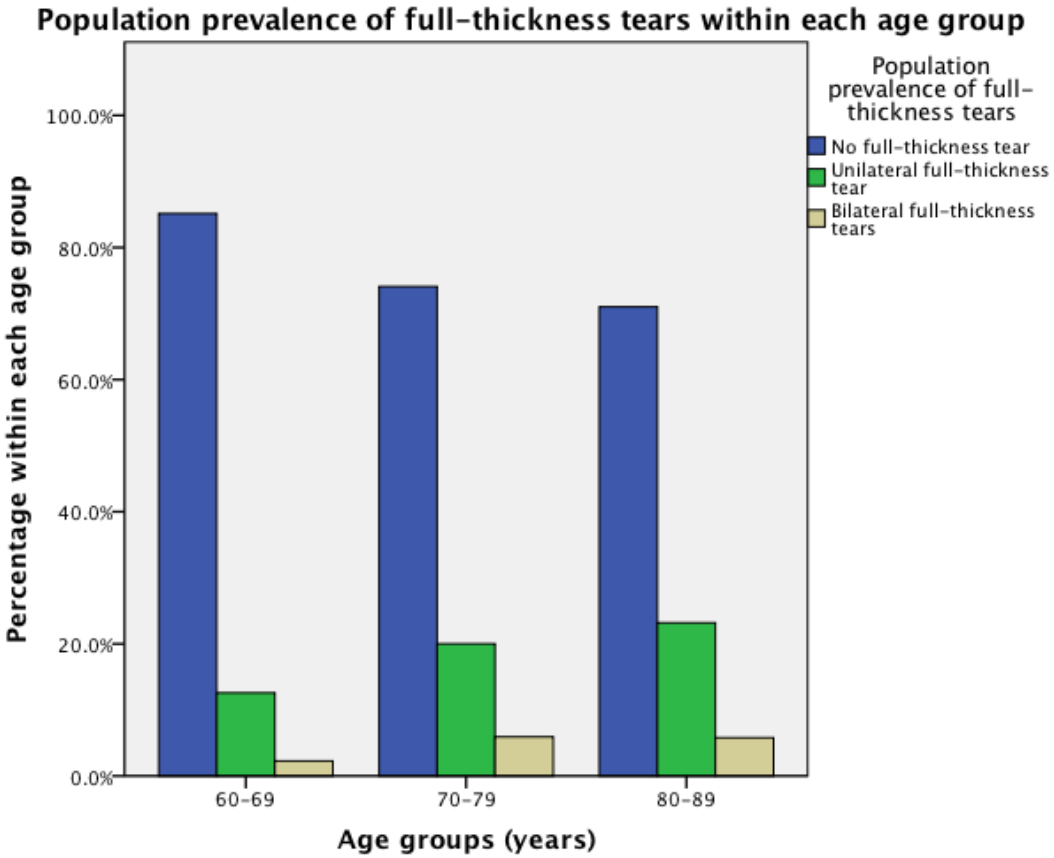
3: POPULATION PREVALENCE

The population prevalence of having at least one full-thickness tear was 22.2% (17.7% unilateral & 4.5% bilateral). For age groups 60-69, 70-79 and 80-89 these were 14.9%; 25.9% and 29% respectively, and bilateral tears 2.3%; 5.9% and 5.8% respectively (see table 14 and graph 25). The difference in population prevalence between age groups was statistically significant (Chi^2 trend p=<0.001). There was no difference in population prevalence between different BMI groups or right and left-handers.

Population prevalence of full-thickness tears according to age decile

	Ages 60-69 (n)	Ages 70-79 (n)	Ages 80-89 (n)	All ages (n)
No full-thickness tear	85.1% (149)	74.1% (163)	71.0% (49)	77.8% (361)
Unilateral full-thickness tear	12.6% (22)	20.0% (44)	23.2% (16)	17.7% (82)
Bilateral full-thickness tear	2.3% (4)	5.9% (13)	5.8% (4)	4.5% (21)
Population prevalence	**14.9% (175)**	**25.9% (220)**	**29.0% (69)**	**22.2% (464)**

Table 14: Population prevalence of full-thickness tears increased significantly with age with an overall prevalence of 22.2%



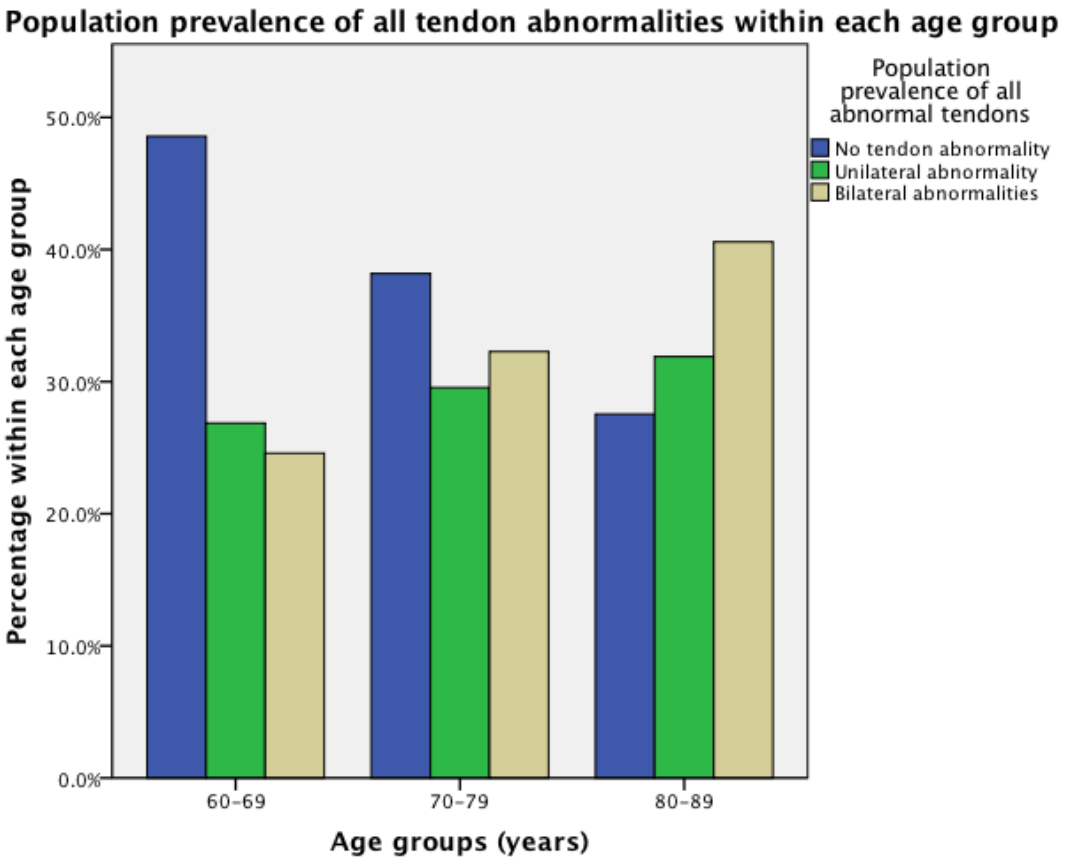
Graph 25: Population prevalence of full-thickness tears increased with age, and those with bilaterally normal shoulders decreased

The population prevalence of having at least one tendinopathy or tear was 59.5% (30.6% bilateral). For age groups 60-69, 70-79 and 80-89 these were 51.5%; 61.8% and 72.5% respectively, and bilateral tears 24.6%; 32.3% and 40.6% respectively (see table 15 and graph 26). The difference in population prevalence between age groups was statistically significant (Chi^2 trend p=<0.001). There was no difference in population prevalence between different BMI groups or right and left-handers.

Population prevalence of rotator cuff abnormalities

	Ages 60-69 (n)	Ages 70-79 (n)	Ages 80-89 (n)	All ages (n)
No abnormality	48.6% (85)	38.2% (84)	27.5% (19)	40.5% (188)
Unilateral abnormality	26.9% (47)	29.5% (65)	31.9% (22)	128.9% (134)
Bilateral abnormality	24.6% (43)	32.3% (71)	40.6% (28)	30.6% (142)
Population prevalence	**51.5% (175)**	**61.8% (220)**	**72.5% (69)**	**59.5% (464)**

Table 15: Population prevalence of rotator cuff abnormalities increased significantly with age, with an overall prevalence of 59.4%



Graph 26: Population prevalence of tendon abnormalities increased with age, and the number with bilaterally normal tendons decreased

5: Discussion

Our study has shown that in women aged between 65 and 84, the overall population prevalence of having at least a unilateral full-thickness tear is 22.2%. This significantly increases with every decile of age, thus for those ages 65-69 it was 14.4%, increasing to 29.0% in those aged 80-84. When considering shoulders individually the relative risk of having a full-thickness tear in the >80 age group was 2.3 times that of the 65-69 age group.

This study is also the first to determine the population prevalence of degenerative tendon abnormalities, including partial-thickness tears. This study

has demonstrated that in women aged between 65 and 84 the general population prevalence of degenerative tendon abnormalities, including partial and full-thickness tears is 59.4%. Again this significantly increased with age, and those aged 65-69 it was 51.5% increasing to 72.5% in those aged 80-84.

Prevalence in all shoulders, not individuals, of either full-thickness tears or any tendon abnormality increased not only with age but also if the shoulder concerned was of the dominant arm. The prevalence of tears in the dominant arm was 15.9% compared to 10.8% in the non-dominant arm (Overall prevalence in all shoulders was 13.4%). The relative risk of full-thickness tear was 1.64 that of the non-dominant arm. BMI and handedness had no effect on prevalence.

Although many studies have attempted to determine the prevalence of full-thickness rotator cuff tears, no study to date has defined this in terms of the general western world population. The major strength of this study is that we have done this using a large general population cohort of women that has been characterised to be representative of the UK general population with regards to weight, height and smoking status, which was not originally set up to study shoulder problems. The selection into this study had no relationship to shoulder pain or shoulder pathology as the initial study was set up to study osteoporosis. This has avoided the selection and surveillance bias that other studies have been subject to by selecting cases from either shoulder clinics or other musculoskeletal clinics. The net effect of such selection bias would be to potentially raise any prevalence significantly.

There are however potential limitations in terms of generalizability with the use of this cohort. Firstly, the cohort only studies women. Previous studies that have included both men and women have not however found any difference in prevalence of rotator cuff tear between men and women and thus this should not influence the prevalence determined in this study. Secondly this cohort contains only those aged over 63, which if generalized to the whole population would give an inflated prevalence due to the known association between rotator cuff tears and increasing age. However this study was designed to investigate degenerate rotator cuff pathology and thus the cohort targets our goal, but our prevalence quoted is the prevalence in the over 60's rather than the whole population.

Although as previously discussed the major strength of this study is using a general population cohort that was not set up to study shoulder problems, there is a potential limitation, as the cohort is in its 20th year of investigation. Having started with 1003 participants and now at the 20th year it containing 516, with 464 being used in this study, it is subject to selection and survival bias. 158 of the cohort have died, 111 were unable to attend the year 20 study, and 218 had either moved away, dropped out or were lost to follow up. The loss of the more elderly group due to death, and the more infirm is unlikely to have any effect on prevalence as firstly we have stratified prevalence by age groups, and secondly there are not any known associations of rotator cuff tears with significant disease, unless the major incapacity was musculoskeletal morbidity. The 218 who were lost to follow up or declined to attend the year 20 visit, are also unlikely to cause selection bias and effect reported prevalence. This is due to

their initial invitation to the study being to investigate osteoporosis, not shoulder pain or pathology. This is a strength of this study as repeat attendance to the study was not influenced by shoulder symptoms, which is a significant selection bias found in all other studies of shoulder pain and structure. Of the 52 who did not undergo a shoulder assessment at the year 20 visit we have demonstrated that demographic profiles between those seen and those who were not are not statistically different, and also as the women were selected at random to attend the visit there proportions with shoulder pain of cuff tear should be randomly distributed.

Other potential limitations of this study are in the methodology. The use of two observers could lead to analytic bias, particularly as ultrasound is a subjective measure. This was overcome in the form of a reproducibility study (see appendix one for full details). Here although a sub-cohort were not examined by both examiners the demographic profiles along with the tear prevalence and symptom association profiles were compared revealing no significant differences. As ultrasound is subjective intra observer reproducibility bias could occur. To demonstrate that this was not significant a small reproducibility study was undertaken which demonstrated no significant differences (see appendix 1 for further details). The final potential limitation is reporting bias. Due to manpower the same assessor performed the entire musculoskeletal assessment and ultrasound for each subject. A protocol was followed for consistency, which involved a verbal consultation, physical examination followed by the ultrasound, thus it was possible for the preceding examination to influence the reported outcome of the shoulder ultrasound. If as the thesis hypothesis suggests, that

rotator cuff tears do have an association with pain, and that the assessor would be more likely to report pathology in the presence of pain, then depending on the prevalence of shoulder pain in the cohort, prevalence of pathology could be affected either way. In this case, as will be shown in the following chapter, the overall prevalence of shoulder symptoms was 25.8%, compared to the prevalence of full-thickness tears being 13.4%, and all abnormalities being 45.1%, thus there would be the potential for both over-reporting and under-reporting of prevalence.

Overall, this study has shown that rotator cuff tears, and degenerate tendon changes are a common age related phenomena. It is the first general population study that is applicable to the UK and like populations. Although this is a cross sectional cohort study, and therefore cannot imply causality, they are more prevalent in the dominant arm and therefore could be associated with use. In order to determine the clinical implications of this epidemiology, we need to explore their relationship with symptoms.

2: The association between rotator cuff tears, shoulder pain and activities of daily living in the general population

1: Background information

The initial study in this thesis has demonstrated that in a normal population of women aged between 65 and 84, the population prevalence of rotator cuff tears was 22.2%, which significantly increased with age. The population prevalence of a tendon abnormality was 59.5%, again increasing significantly with age.

However less well established in the existing literature, particularly in relation to the general population, is the relationship between pathology and symptoms. The clinical manifestations of rotator cuff tears are varied, as is highlighted by the number of studies focusing on asymptomatic tears, and this lack of understanding of the relationship between pathology and symptoms may contribute to the literature's lack of consensus regarding the management of rotator cuff tears. Understanding the natural progression of tears and the relationship with symptoms will help to understand and develop more effective treatment strategies.

Previous studies have demonstrated that clinical manifestations of rotator cuff tears can be varied, including between the two arms of an individual. Yamaguchi et al. performed bilateral ultrasounds in patients with painful rotator cuff tears,

finding a 35-50% increase in chance of a contralateral asymptomatic rotator cuff tear, increasing with age. The only difference between the shoulders was an increase in tear size in the symptomatic shoulder (66). Attempts have been made to follow up these asymptomatic cohorts to better understand the natural history of full-thickness tears and pain. Yamaguchi showed 51% became symptomatic over a mean 2.8 year period with 50% of those progressing in size (71). In a similar cohort Mall et al. revealed that those who became symptomatic initially had larger tears than those who remained asymptomatic, and furthermore, 23% had progressed in size compared to only 4% in those remaining asymptomatic (67). This study also looked at gleno-humeral migration and found this not to be significantly different between groups. Moosmayer et al. had similar results in prospective follow up study of 195 asymptomatic full-thickness tears. 36% became symptomatic over a 3 year period (75). Initially there was no difference in tear size between those that became symptomatic and those remaining symptomatic, but 3 years on there was significant difference. These studies all point towards tear size being a contributing factor towards the development of symptoms. However, the cohorts have all been derived from either orthopaedic clinics or from cohorts where data from the contralateral shoulder of a painful full-thickness tear was used, thus do not necessarily represent the general population.

One mountain population cohort study has been performed to attempt to identify factors associated with painful full-thickness tears. Here 34.6% of all full-thickness tears were painful, and the factors associated with pain included if

being the dominant arm, or the presence of impingement signs. No tear measurement data, or tendinopathy data was evaluated (76).

The severity of pain has also been evaluated. Harris et al. found in a cohort of symptomatic full-thickness tears undergoing conservative treatment, that the increase in severity of pain was associated with the male sex, scapular dysfunction, supraspinatus atrophy, and multi tendon tears. They did not find association with isolated tendon tear size, age or extent of humeral head migration (77).

To date, no studies have yet explored the associations between rotator cuff tears, pain and functional loss in a general population cohort, which would provide baseline epidemiology to the natural history of rotator cuff tears. Also no studies have looked either at the relationship of pathology and symptoms with degenerate tendons that are not torn, as well as partial-thickness tears, or the possible change in severity of symptoms across different stages of pathology. Fundamentally, this study will demonstrate the likelihood and severity of symptoms in the general population across all stages of rotator cuff tendinopathy and will greatly contribute to the overall understanding of the natural history of rotator cuff tears.

2: Aims and objectives

The primary aim is to test the hypothesis that rotator cuff tears are associated with pain and qualitative functional loss, specifically:

1. Are different stages of rotator cuff tear associated with an increased prevalence of pain and qualitative functional loss?
2. Does the severity of the pathology correlate with the severity of symptoms?

The secondary aim is to look from a clinical perspective, at the proportion of those with symptomatic rotator cuff pathology who have sought medical advice.

3: Patients and methods

PATIENTS

The study participants were selected from the Chingford Study, a well described prospective population-based longitudinal study of osteoarthritis and osteoporosis, comprising initially of 1003 women, derived from the register of a large general practice in Chingford, North London (see section 1.1). The women aged 44-67 years at baseline are representative of women in the UK general population with respect to weight, height, and smoking characteristics. At year 20 of the study 463 underwent a shoulder musculoskeletal assessment including the Oxford shoulder score, and bilateral shoulder ultrasound (Of the original 1003, 158 women had died, 111 were unable to attend, 218 had either moved away, dropped out or were lost to follow up, and 52 attended the year 20 visit but did not have a musculoskeletal assessment due to absence of the assessor, and one shoulder data sheet was not completed). The local ethics committee approved the study and consent was obtained from each woman (Outer North

East London Research Ethics Committee (formerly Barking and Havering and Waltham Forest RECs), LREC (R&WF) 96).

METHODS

For each participant age, height measured in cm (to the nearest 0.1cm) in a standing position with shoes removed, using a wall-mounted stadiometer (Leicester height measure, Seca). Weight was measured in kg (to the nearest 0.1kg) by electronic scales with shoes removed were recorded. A fixed SOPP (standard operating procedure protocol) was followed for each musculoskeletal assessment. This was followed by ultrasound assessment of the shoulders (left then right).

Symptom assessment

Symptoms have been defined using the validated Oxford shoulder score (30,31), which is a self-administered questionnaire validated for the study of shoulder pain. This was completed by the participant prior to attending the study and was reviewed with the clinician during the assessment to reduce error. Binary symptoms were defined using a dichotomised OSS score where any non-perfect OSS was classified as symptomatic. Where questions are pain specific, the four pain specific questions of the OSS were used as a sub-scale and any non-perfect score in the pain subset was classified as symptomatic, whilst perfect scores were asymptomatic. In those with symptoms the full OSS scale was used to define symptom severity. Participants were also asked whether they had ever sought medical advice regarding shoulder pain.

For justification and validation of the use of the OSS as the primary outcome measure see section 1:3.

Subsequent to our initial analysis using a single point change in the OSS to define symptoms we have also performed a comparative analysis using a 3-point change in the OSS (or subset of) to define symptoms. Thus any score of 45 or less (or 13 for the pain subset, and 29 for the function subset) was classified as symptomatic, and scores greater were defined as asymptomatic. This was performed to reflect the clinical relevance of symptoms. For justification of using the 3-point change see section 1.3.4.

Ultrasound examination

The ultrasound examination of the 463 women was performed using the GE voluson i portable ultrasound machine with a 10-16MHz linear probe and is detailed in appendix 2. A fixed protocol was derived according to the recommendations of the Nuffield Orthopaedic Centre Musculoskeletal Radiology Department, and results were recorded on a predetermined data entry sheet. Tendons were classified into one of 4 working groups based upon ultrasound measurements:

1. Normal tendons
2. Abnormal tendons and partial thickness tears
3. Single tendon full-thickness tears (0-2.5cm)
4. Multi-tendon full-thickness tears (>2.5cm)

Justification of the development and use of the above tendon classification, along with validation studies is detailed in section 1.2.

STATISTICAL METHODS

Age, BMI, arm dominance, and symptom presence were compared across the 4-tendinopathy groups. Wilcoxon rank sum test, one-way ANOVA, and chi-squared tests were used for non-normal, normal and categorical data respectively.

To answer whether different stages of rotator cuff tear associated with an increased prevalence of pain and qualitative functional loss, a dichotomised OSS scale was used separating shoulder into binary symptomatic and asymptomatic groups. This was firstly due to the requirement of a binary outcome measure and secondly due to the highly positively skewed OSS data, which could not be transformed to normal when using the full cohort. A Chi^2 test was used to determine difference between tendinopathy groups. Multivariable binary logistic regression was used to adjust for the potential confounders age, and hand dominance all of which were determined a priori. Whether it was the dominant or non-dominant arm had no effect on the likelihood of symptoms, therefor there was no requirement for a multi-level model. Thus all 926 shoulders from 463 individuals were included in the same single-level model. However to understand the relationship between the individual, structural pathology and symptoms tables were drawn up comparing those with bilateral or unilateral structural pathology, or lack of, with their symptoms.

To determine whether the severity of symptoms was affected by the severity of pathology all asymptomatic shoulders were removed. Logarithmic transformation of the inverse OSS then created a normal distribution. The severity of symptoms in these symptomatic shoulders was compared across tendinopathy groups using a 1-way ANOVA. Multivariable linear regression was used to adjust for potential confounders age, and hand dominance all of which were determined a priori. As above whether data was obtained from the dominant or non-dominant arm did not affect symptom association, thus all shoulders were included in a single-level regression model.

A justification of this statistical approach, with details of data transformations and statistical modelling options is found in appendix 4.

All statistics were performed using IBM SPSS Statistics version 22 (IBM, Armonk, NY, USA).

4: Results

1: Symptom prevalence

Descriptive statistics

463 individuals (926 shoulders) were included in the study. Age was not normally distributed and the median age was 71 (range 65-84). BMI was normally distributed with a mean of 27.8. Of the 516 individuals attending the year 20 visit, the median age was 71 (range 63-85), and mean BMI was 27.8,

there was no statistical difference in age or BMI between groups (p=0.222/p=0.136).

The demographics of the 926 shoulders are summarised in table 16. 508 were diagnosed as normal on ultrasound, 294 abnormal or partial-thickness tears, and 85 full-thickness tears (0-2.5cm), and 39 full-thickness tears (>2.5cm). The distribution of age across each tendinopathy group was significantly different (Kruskal Wallis Test p<0.001). There was no statistical difference in BMI for each tendinopathy group (One way ANOVA p=0.080). 420 (90.5%) individuals were right-handed, and 44 (9.5%) were left-handed. There was no difference in right and left-handed individuals in each group (Chi^2 test p=0.771), however there was a statistical difference in dominant and non-dominant arms in each tendinopathy group (Chi^2 test p=0.033), with significantly more non-dominant arms in the normal tendon group and more dominant arms in each of the full-thickness tear groups.

Demographics of different tendinopathy groups

	N	Median age	Mean BMI	Dominant arm (%)
Normal	508	70	27.5	46.1%
Abnormal/Partial tear	294	73	28.0	52.7%
Full-thickness tears 0-2.5cm	85	74	27.9	58.8%
Full-thickness tears >2.5cm	39	74	29.6	61.5%
All	926	71	27.8	50%

Table 16: Age and the proportion of dominant arms increased with increasing tendinopathy, there was no difference in BMI between groups

Primary outcome measure: Dichotomised OSS

289 (31.2%) of shoulders were defined as symptomatic according to the binary OSS, the demographics of which are shown in table 17. There was no difference in age, BMI, or which hand the data came from. The number of symptomatic tendons for each tendinopathy group was as follows: Normal tendons, 116/508 (22.8%); abnormal/partial tears, 109/294 (37.1%); full-thickness tears 0-2.5cm, 35/85 (41.2%); full-thickness tears >2.5cm, 29/39 (74.4%), as shown in table 18. The proportion with symptoms was statistically significant between tendinopathy groups, with a greater proportion of symptoms as tear stage severity increased (Chi^2 linear association $p<0.001$). This is shown in graph 27.

Demographics of symptomatic and asymptomatic shoulders

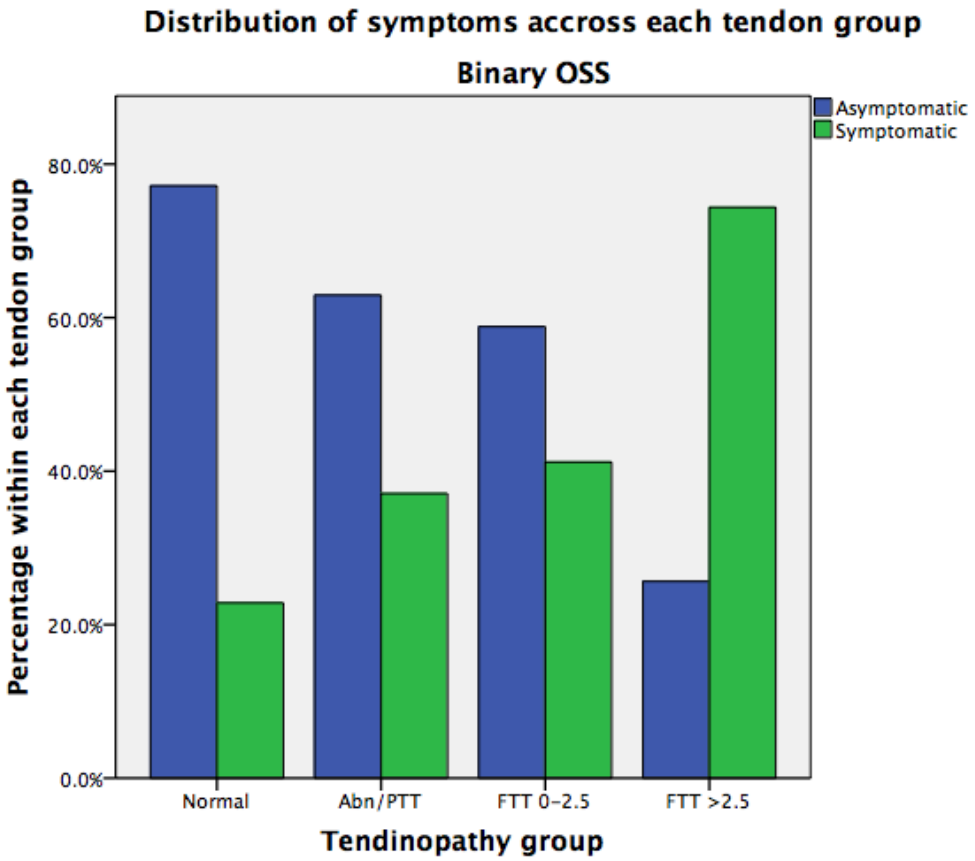
	N	Median age	Mean BMI	Dominant arm (%)
Symptomatic	289	72	28.5	52.6%
Asymptomatic	637	71	27.4	47.4%
All	926	71	27.8	50%

Table 17: Symptoms were more prevalent in the dominant arm

Distribution of symptoms across each tendinopathy group

	N	Symptomatic	% with symptoms
Normal	508	116	22.8%
Abnormal/Partial tear	294	109	37.1%
Full-thickness tears 0-2.5cm	85	35	41.2%
Full-thickness tears >2.5cm	39	29	74.4%
All	926	289	31.2%

Table 18: Symptom prevalence increased as tear stage severity increased



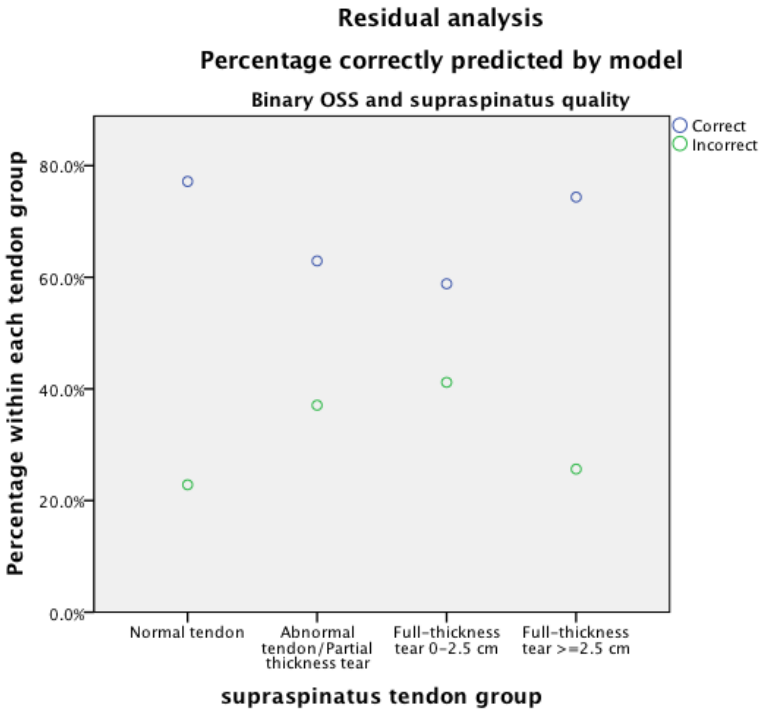
Graph 27: Symptom prevalence increased as tear stage severity increased

Binary logistic regression analysis after adjustment for BMI, age, and whether it was the dominant or non-dominant arm (the latter two of which are known to be associated with rotator cuff pathology) gave the following relative risks (RR) of having symptoms compared to the normal tendon: Abnormal/Partial tears 1.969; full-thickness tears 0-2.5cm 2.203; and full-thickness tears >2.5cm 4.718. All were significant at $p<0.001$ with the model correctly predicting 71% of symptom outcomes correctly. This is shown in table 19, and the residual analysis demonstrated in graph 28. There was no confounding effect or interactions with age, BMI, or whether it was the dominant or non-dominant arm, as is illustrated in the below graphs 29-31.

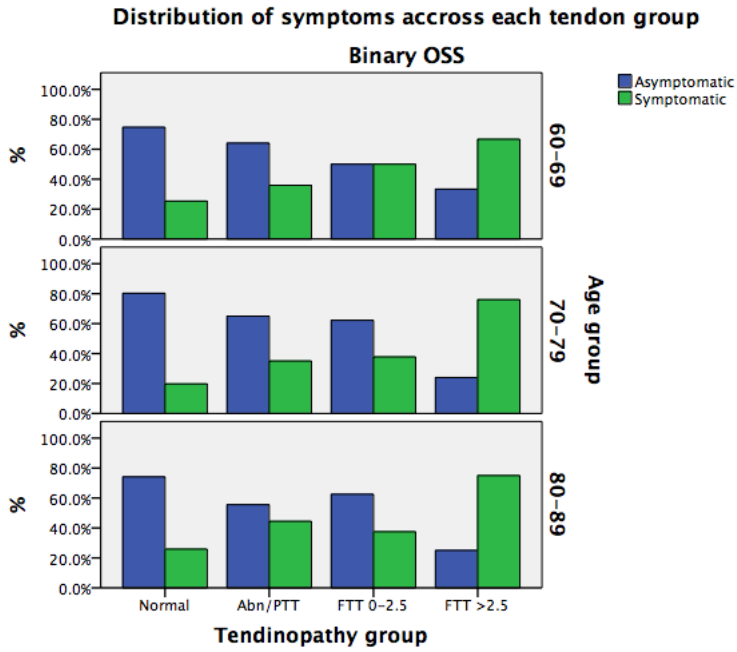
Regression analysis of binary OSS and supraspinatus quality

	OR	RR	Sig.	95% C.I.for OR	
				Lower	Upper
Constant (Normal tendons)	.296		<0.001		
Tendon Classification			<0.001		
Abnormal / partial tears	1.991	1.969	<0.001	1.454	2.727
Full-thickness tears ≤2.5cm	2.366	2.203	<0.001	1.465	3.819
Full-thickness tears >2.5cm	9.800	4.718	<0.001	4.638	20.705

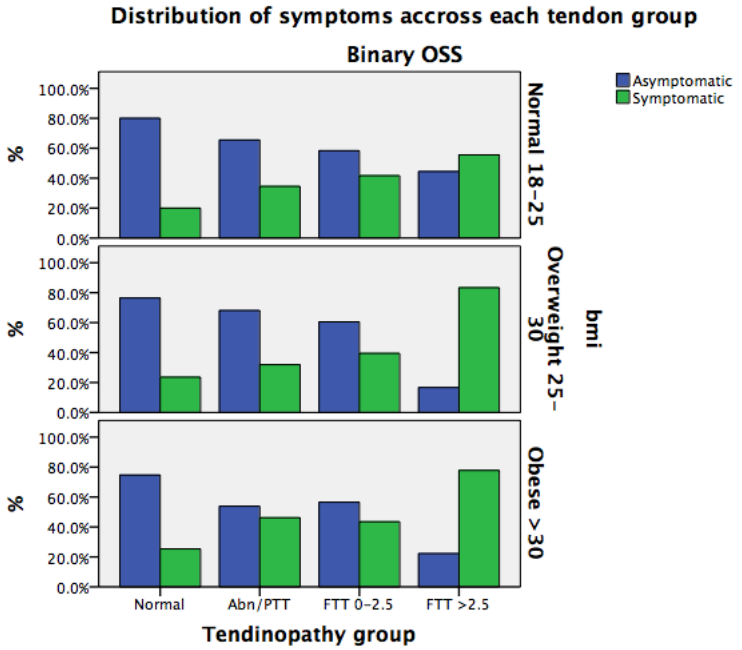
Table 19: The relative risk of symptoms increased with tear stage severity



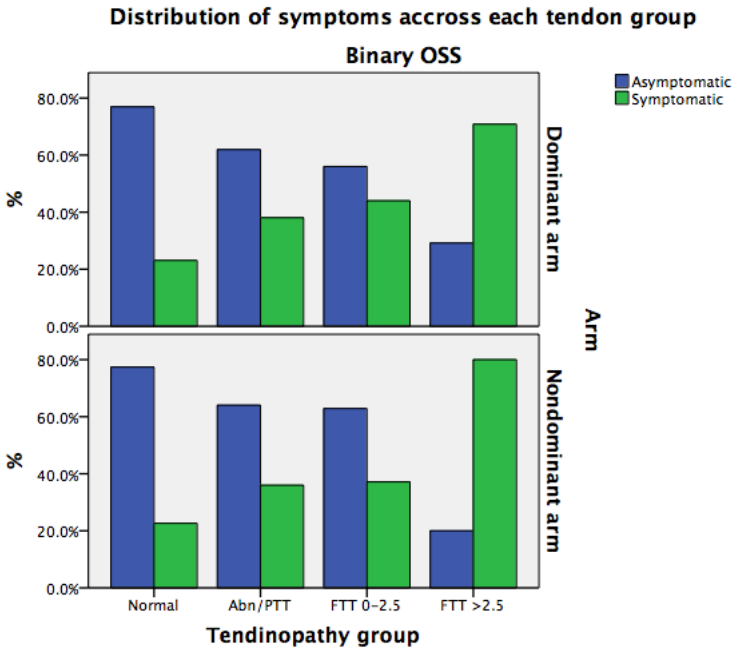
Graph 28: Residual analysis of regression model indicating that the model predicted the correct outcome in 71% of cases



Graph 29: There was no difference is symptoms across different age groups



Graph 30: There was no difference in symptoms across different BMI groups



Graph 31: There were more symptomatic shoulders in the dominant arm group

Primary outcome measure dichotomised OSS (3 point change version)

The above analysis was re-performed where symptoms were defined using a 3-point change in the OSS rather than a single point as above. The results are shown in tables 20-21.

Distribution of symptoms across different tendinopathy groups using a 3-point change in the OSS

	N	Symptomatic	% with symptoms
Normal	508	87	17.1%
Abnormal/Partial tear	294	84	28.6%
Full-thickness tears 0-2.5cm	85	29	34.1%
Full-thickness tears >2.5cm	39	26	66.7%
All	926	226	24.4%

Table 20: Symptom likelihood increased with increasing tear stage severity

Regression analysis of binary OSS and supraspinatus quality

	OR	RR	Sig.	95% C.I.for OR	
				Lower	Upper
Constant (Normal tendons)	.207		.000		
Tendon Classification			.000		
Abnormal / partial tears	1.936	1.793	.000	1.374	2.726
Full-thickness tears ≤2.5cm	2.506	2.098	.000	1.513	4.150
Full-thickness tears >2.5cm	9.678	3.924	.000	4.784	19.580

Table 21: Relative risks of symptoms increased with increasing tear stage severity

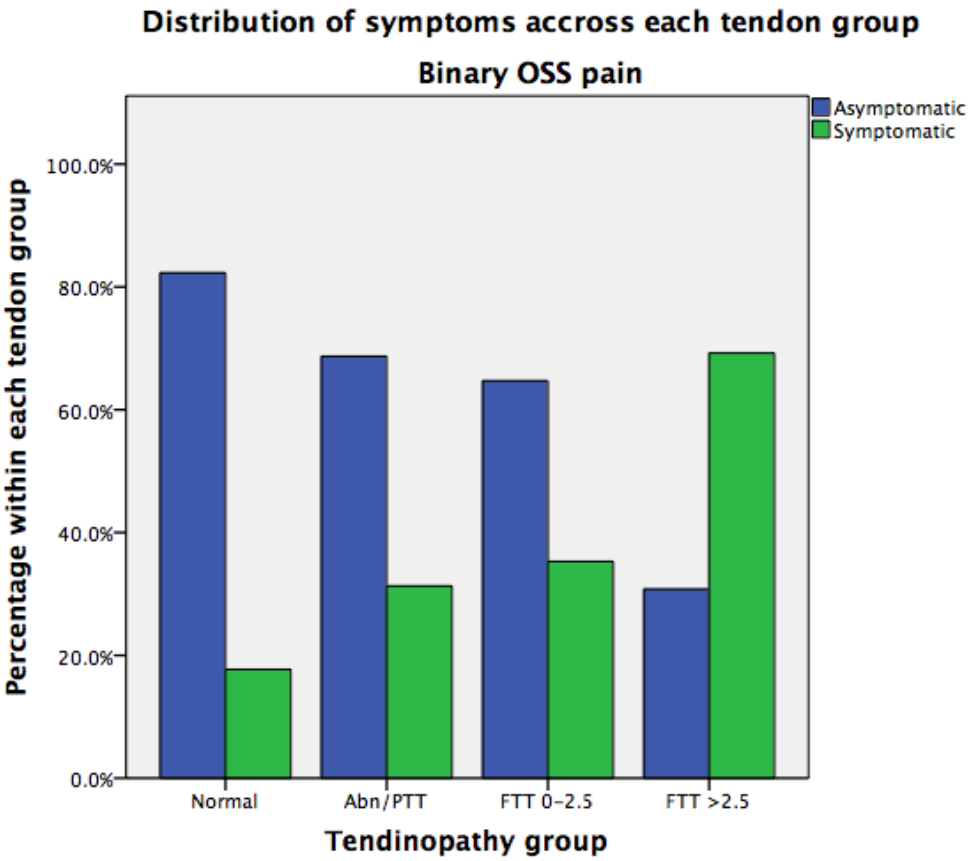
Secondary outcome measures: dichotomised OSS (pain) and OSS (function)

239 (25.8%) of shoulders were classified as painful according to the binary OSS pain scale. 238 (25.8%) of shoulders were classified as having qualitative functional loss according to the binary OSS function scale. The number of symptomatic tendons for each tendinopathy group are shown in table 14. The proportion with symptoms was statistically significant between tendinopathy groups in both the pain and function subscales, with both pain and function being more prevalent as pathology increases (Chi^2 linear association $p<0.001$). This is shown in table 22 and graphs 32-33.

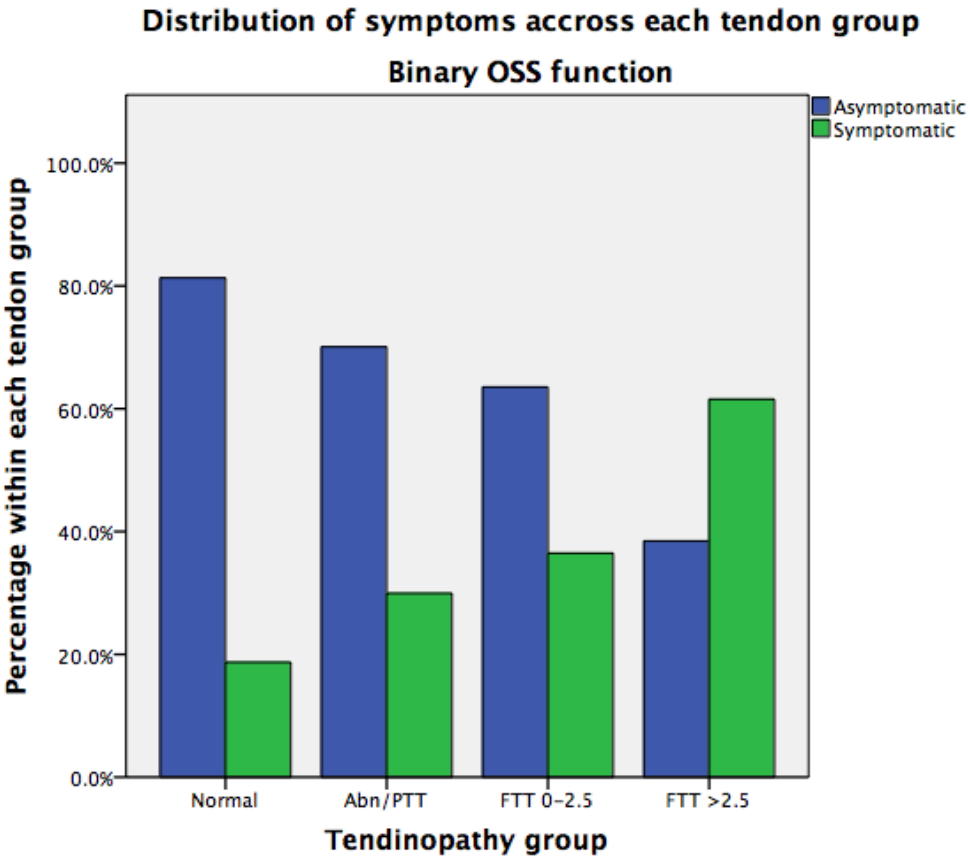
Pain and function distribution across different tendinopathy groups

	N	Pain	% with pain	Function loss	% with function loss
Normal	508	90	17.7%	95	18.7%
Abnormal/Partial tear	294	92	31.3%	88	29.9%
Full-thickness tears 0-2.5cm	85	30	35.3%	31	36.5%
Full-thickness tears >2.5cm	39	27	69.2%	24	61.5%
All	926	239	25.8%	238	25.7%

Table 22: Symptoms, whether pain of function increased with increasing tear stage severity



Graph 32: The proportion with pain increased according to tear stage severity



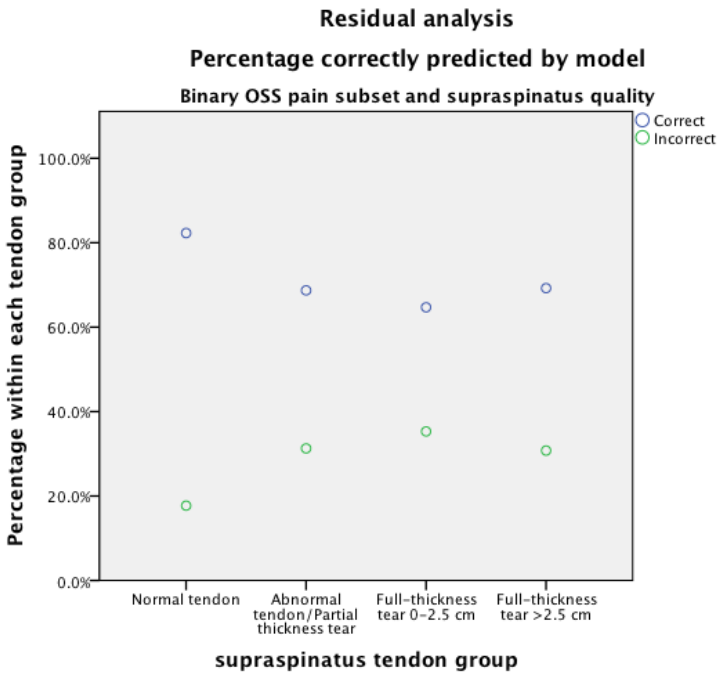
Graph 33: The proportion with functional deficits increased according to tear stage severity

The following tables and graphs (tables 23-24, graphs 16 and 17) show the binary logistic regression analysis and residual analysis after adjustment for age, BMI, and hand dominance for both the pain and function subscales. There was no confounding effect or interactions with age, BMI, or whether it was the dominant or non-dominant arm. The increased RRs were all significant at p<0.001 with the model correctly predicting 75.8% and 75.3% of symptom outcomes correctly for the pain and function scales respectively.

Regression analysis of binary OSS (pain) and supraspinatus quality

	OR	RR	Sig.	95% C.I.for OR	
				Lower	Upper
Constant (normal tendons)	.215		<0.001		
Tendon classification			<0.001		
Abnormal / partial tears	2.115	2.063	<0.001	1.512	2.959
Full-thickness tears ≤2.5cm	2.533	2.349	<0.001	1.537	4.176
Full-thickness tears >2.5cm	10.450	5.499	<0.001	5.101	21.407

Table 23: Relative risk of pain increased with tear stage severity

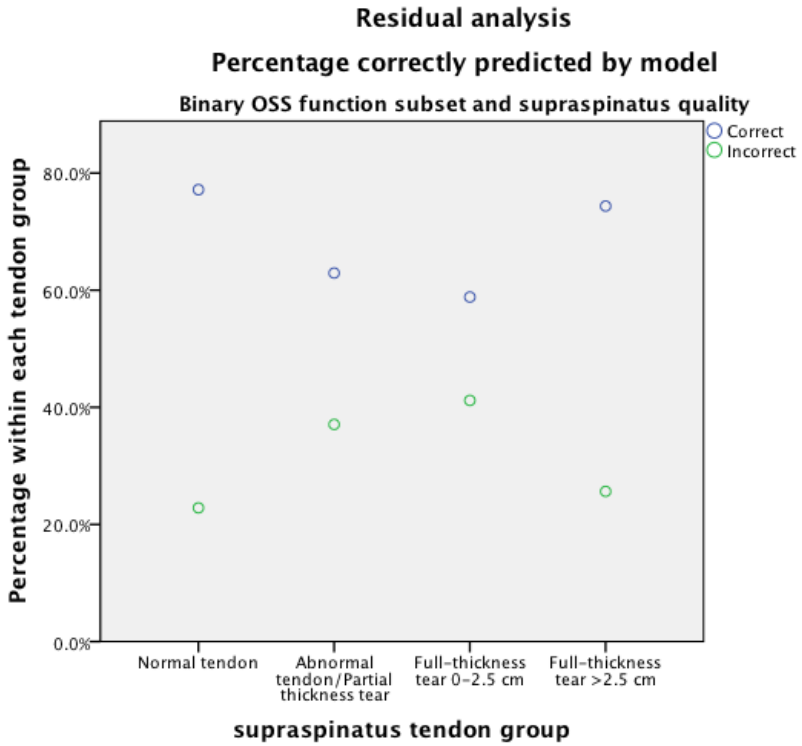


Graph 34: Residual analysis of regression model showing the model predicted 75.8% of cases correctly

Regression analysis of binary OSS (function) and supraspinatus quality

	OR	RR	Sig.	95% C.I.for OR	
				Lower	Upper
Constant (normal tendons)	.230		.000		
Tendon classification			<0.001		
Abnormal / partial tears	1.857	1.878	<0.001	1.329	2.595
Full-thickness tears ≤2.5cm	2.496	2.316	<0.001	1.522	4.093
Full-thickness tears >2.5cm	6.956	4.396	<0.001	3.515	13.765

Table 24: Relative risk of functional deficit increased with tear stage severity



Graph 35: Residual analysis of regression model showing the model predicted 75.8% of cases correctly

Secondary outcome measures: Dichotomised OSS (pain) and OSS (function) (3-point change version)

The following analysis was re-performed where symptoms were defined using a 3-point change in the OSS sub-score rather than a single point as above. The results are shown in tables 25-27.

Distribution of symptoms across different tendinopathy groups

	N	Pain	% with pain	Function loss	% with function loss
Normal	508	76	15.0%	51	10.0%
Abnormal/Partial tear	294	78	26.5%	54	18.4%
Full-thickness tears 0-2.5cm	85	23	27.1%	20	23.5%
Full-thickness tears >2.5cm	39	26	66.7%	20	51.3%
All	926	203	25.8%	145	15.7%

Table 25: Pain and functional loss both became more likely with increasing tear stage severity

Regression analysis of binary OSS (pain) and supraspinatus quality

	OR	RR	Sig.	95% C.I.for OR	
				Lower	Upper
Constant (normal tendons)	.178		.000		
Tendon classification			.000		
Abnormal / partial tears	2.096	1.909	.000	1.461	3.006
Full-thickness tears ≤2.5cm	2.227	1.986	.004	1.291	3.841
Full-thickness tears >2.5cm	11.802	4.526	.000	5.724	24.332

Table 26: Relative risk of pain increased with tear stage severity

Regression analysis of binary OSS (function) and supraspinatus quality

	OR	RR	Sig.	95% C.I.for OR	
				Lower	Upper
Constant (normal tendons)	.090		.000		
Tendon classification			.000		
Abnormal / partial tears	2.005	1.888	.001	1.320	3.044
Full-thickness tears ≤2.5cm	2.818	2.442	.001	1.568	5.066
Full-thickness tears >2.5cm	9.329	5.114	.000	4.596	18.939

Table 27: Relative risk of functional deficit increased with tear stage severity

2. Severity of symptoms

Descriptive statistics

289 shoulders were symptomatic according to the dichotomized OSS and included in the analysis, which are summarised in table 28. Of these 116 were diagnosed as normal on ultrasound, 109 abnormal or partial-thickness tears, and 35 full-thickness tears (0-2.5cm), and 29 full-thickness tears (>2.5cm). Age was not normally distributed. The median age was 71 (range 65-84). The distribution of age across each tendinopathy group was significantly different (Kruskal-Wallis Test p=0.043), driven by the normal tendon group. BMI was normally distributed with a mean of 27.8 and there was no statistical difference in BMI for each tendinopathy group (One way ANOVA p=0.346). 260 (90.0%) were right-handed, and 29 (10.0%) were left-handed. There was no difference in right and left-handed individuals in each group (Chi^2 test p=0.110). There was no statistical difference in dominant and non-dominant arms in each tendinopathy group (Chi^2 test p=0.295).

Demographics of each symptomatic tendinopathy group

	N	Median age	Mean BMI	Dominant arm (%)
Normal	116	70	28.3	46.6%
Abnormal/Partial tear	109	73	28.4	54.1%
Full-thickness tears 0-2.5cm	35	72	28.1	62.9%
Full-thickness tears >2.5cm	29	73	30.3	58.6%
All	289	71	28.5	50%

Table 28: Both age and the proportion of dominant arms increased with increasing tear stage severity

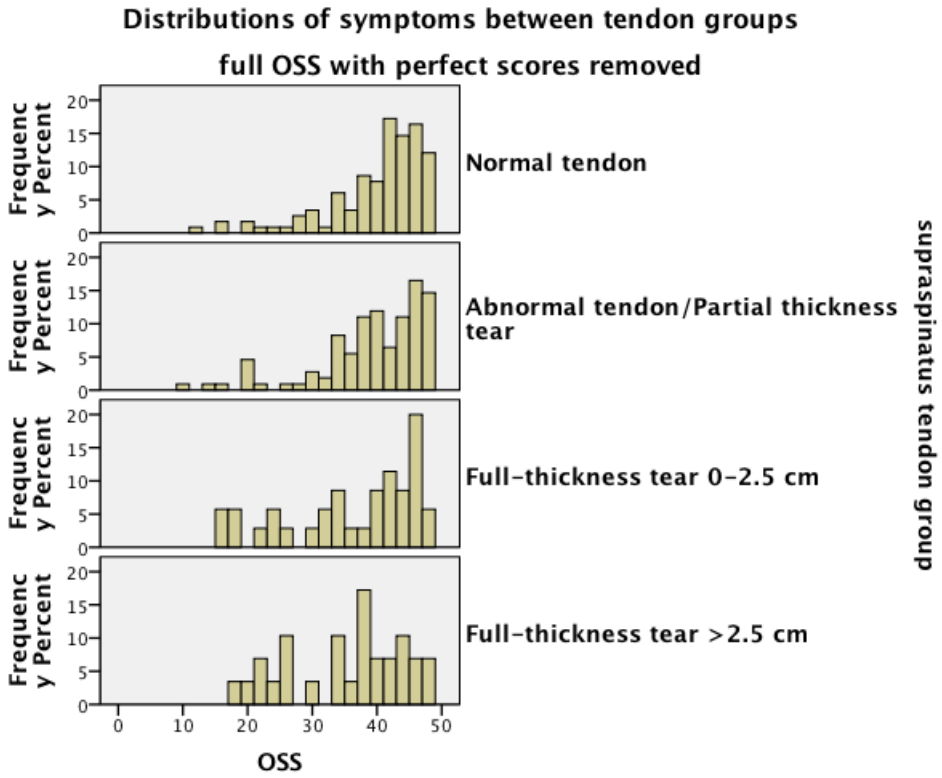
Primary outcome measure OSS

After a natural logarithmic transformation of the inversed OSS score to achieve normality (see appendix 1, section 3.1 which describes the needs for data transformation), the mean OSS was 41.8. For normal tendons this was 42.5, Abnormal tendons 42.1, full-thickness tears 0-2.5cm 40.2, and full-thickness tears >2.5cm 38.4, see table 29 and graphs 36-37. Comparing untransformed data there was a difference in OSS scores between groups (Kruskal-Wallis p=0.020) and transformed groups (1 way ANOVA p=0.030), the difference here driven by the two full-thickness tear groups.

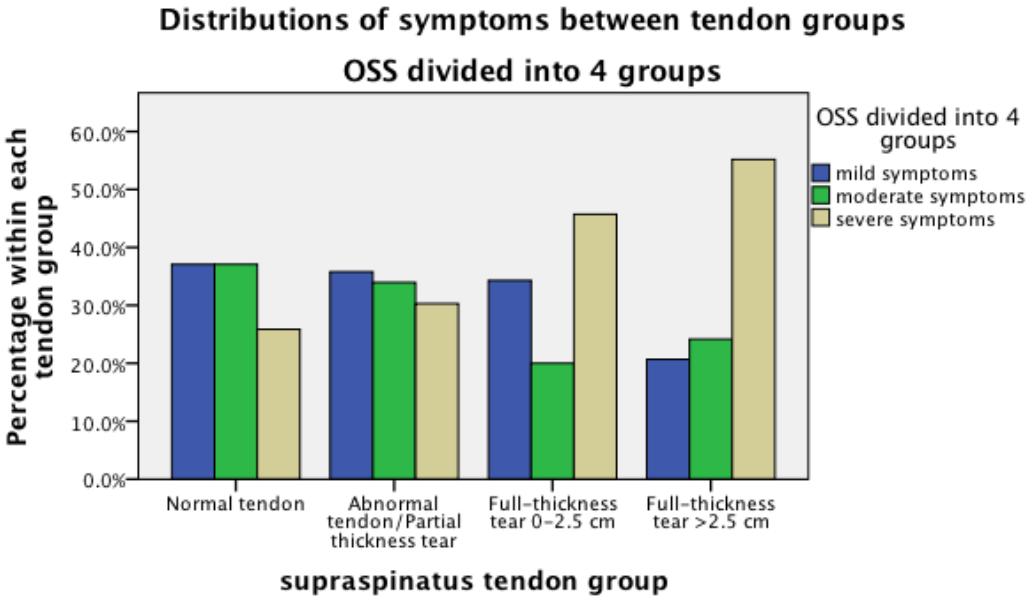
Mean Oxford shoulder scores for each tendinopathy group

	N	Mean OSS	Mean OSS Pain	Mean OSS function
Normal	116	42.5	11.1	29.2
Abnormal/Partial tear	109	42.1	10.9	28.7
Full-thickness tears 0-2.5cm	35	40.2	10.1	27.4
Full-thickness tears >2.5cm	29	38.4	9.1	26.3
All	289	41.8	10.7	28.6

Table 29: All scores decreased with increasing tear stage severity, but this was more apparent in the full OSS and the function subscales



Graph 36: The mean OSS was lower in the FTT >2.5cm group



Graph 37: The severity of symptoms seemed to increase mainly at the FTT>2.5cm group

Linear regression analysis after adjustment for age, (no interactions or confounders identified), showed that the only significant difference in OSS scores was between normal tendons (mean OSS 42.5) and full-thickness tears <2.5cm

(OSS 38.46), p=0.009, power 0.75 (overall model p=0.007, power 0.892) (see table 30).

Regression co-efficients for the OSS in different tendinopathy groups

	Unstandardized Coefficients				95.0% Confidence Interval for B	
	OSS	B	Sig.	Power	Lower Bound	Upper Bound
Normal tendons (Constant)	42.41	1.721	<0.001	1.000	1.499	1.943
Abnormal/Partial tears	42.18	.095	.479	.109	-.168	.358
Full-thickness tears (≤2.5cm)	40.02	.356	.065	.454	-.138	.734
Full-thickness tears (>2.5cm)	38.29	.552	.009	.744	.173	.965

Table 30: The mean OSS only significantly decreased in the FTT>2.5cm group compared to normal symptomatic tendons

Secondary outcome measure OSS pain and OSS function

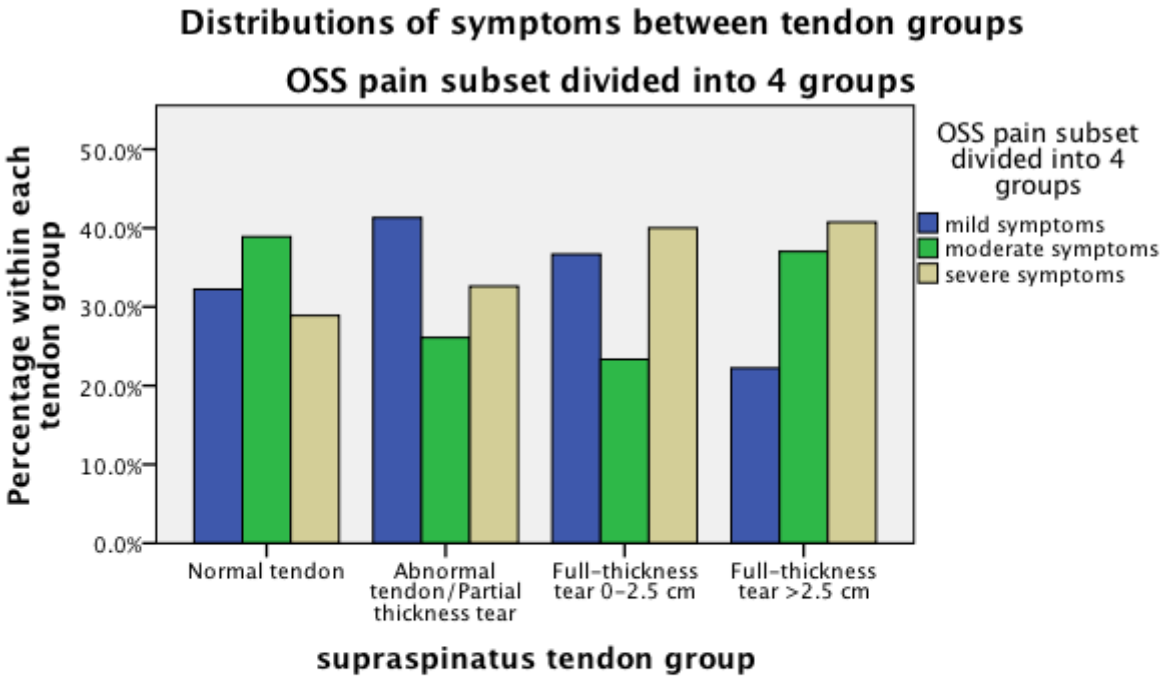
The inversed OSS function scale underwent a natural logarithmic transformation to achieve normality, whilst the OSS pain score was used untransformed. The mean OSS pain score was 10.8/16. For normal tendons this was 10.8, Abnormal tendons 11.1, full-thickness tears 0-2.5cm 10.7, and full-thickness tears >2.5cm 9.5 (see table 31 and graphs 38-39). There was no difference in OSS pain scores between groups (Kruskal-Wallis p=0.416) (1 way ANOVA p=0.088).

The mean OSS function score was 28.6/32. For normal tendons this was 29.2, Abnormal tendons 28.7, full-thickness tears 0-2.5cm 27.4, and full-thickness tears >2.5cm 26.3 (see table 31 and graphs 40-41). Analysis of the untransformed and transformed OSS function scores suggests there is a difference between scores in each tendon group (Kruskal-Wallis p=0.015)(1 way ANOVA p=0.003), which is driven predominantly by the two full-thickness tear groups.

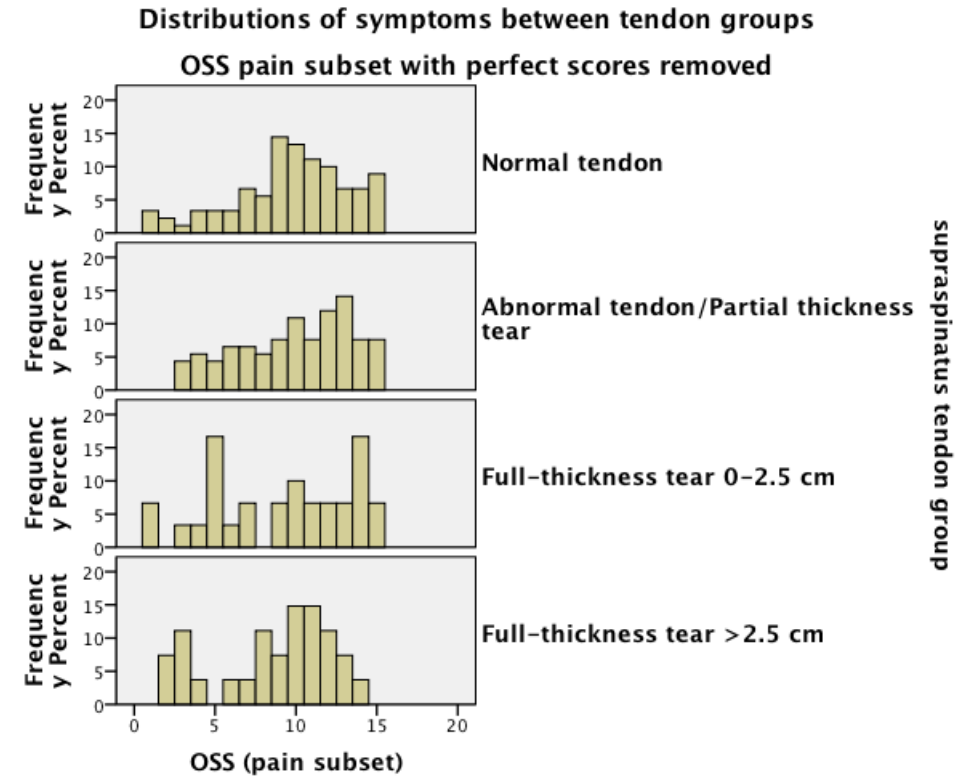
Mean OSS (pain and function) scores in different symptomatic tendinopathy groups

	N	Mean OSS	Mean OSS Pain	Mean OSS function
Normal	116	42.5	10.8	29.2
Abnormal/Partial tear	109	42.1	11.1	28.7
Full-thickness tears 0-2.5cm	35	40.2	10.7	27.4
Full-thickness tears >2.5cm	29	38.4	9.5	26.3
All	289	41.8	10.8	28.6

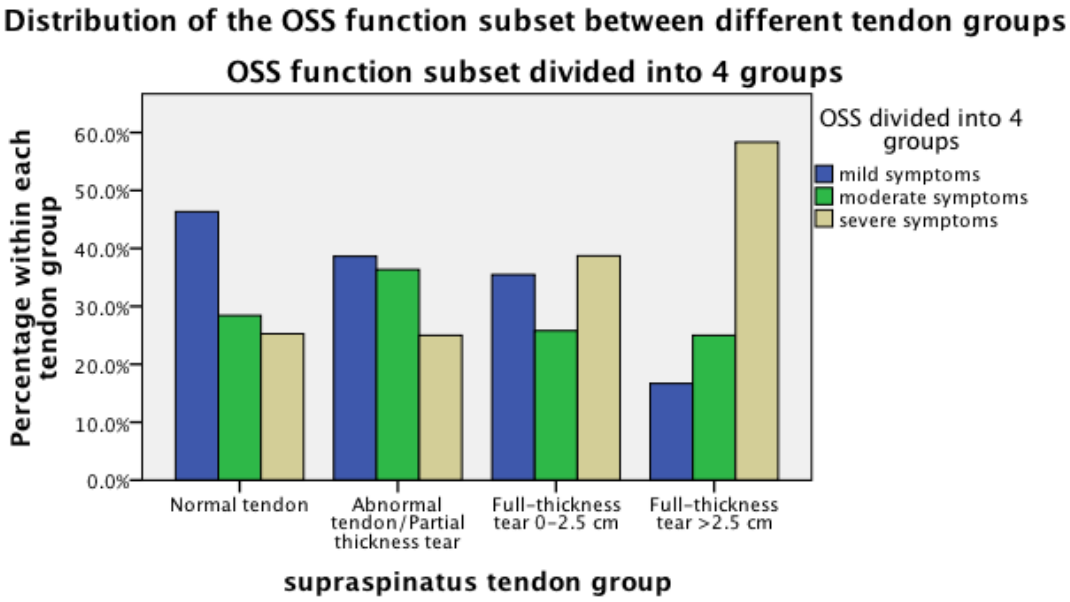
Table 31: All scores decreased with increasing tear stage severity, but this was more apparent in the full OSS and the function subscales



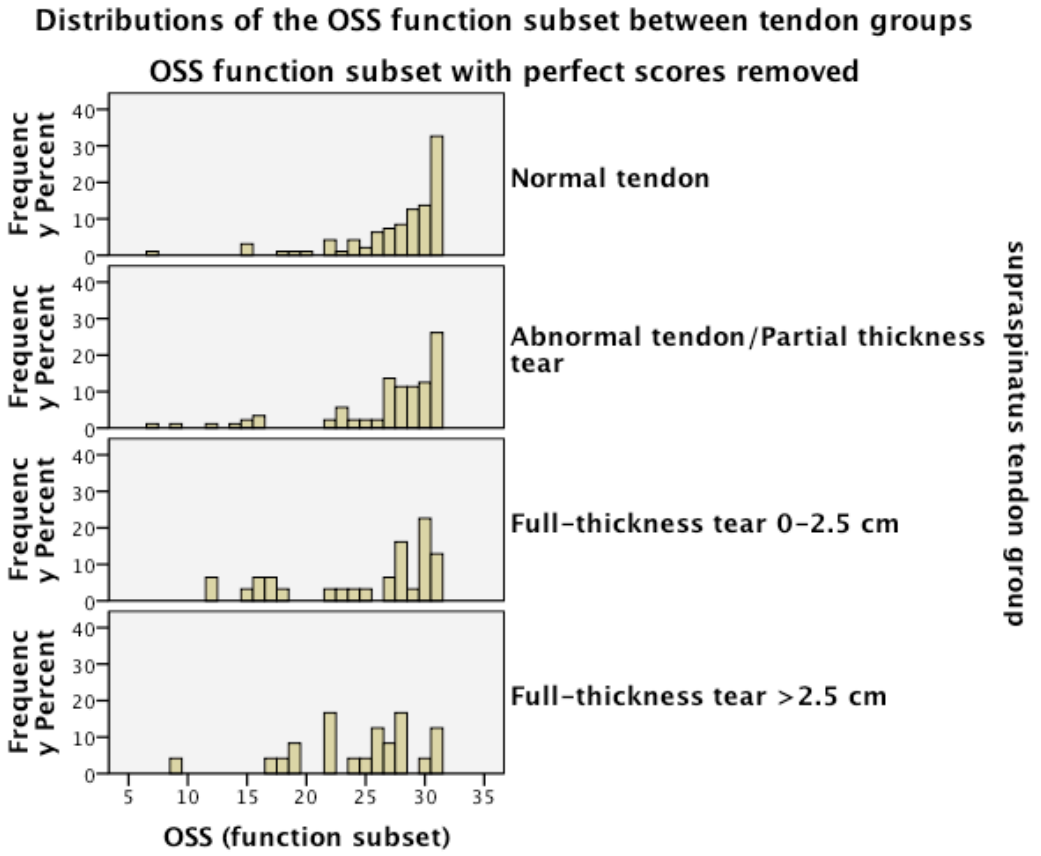
Graph 38: Pain severity did not obviously increase with increasing tear stage severity



Graph 39: Pain severity did not obviously increase with increasing tear stage severity



Graph 40: Symptom severity appears to increase at the FTT>2.5cm group



Graph 41: Symptom severity appears to increase at the FTT>2.5cm group

Linear regression analysis after adjustment for age, BMI, and hand dominance (no interactions or confounders identified), showed that no significant difference in OSS pain scores was between any groups (p=0.335). Linear regression analysis after adjustment for age, BMI, and hand dominance (no interactions or confounders identified), showed a significant difference in OSS function scores between normal tendons (mean OSS function 29.2) and both small and large full-thickness tears (OSS function 27.44 and 26.32 respectively), (p=0.012 and p=0.001 respectively) (overall model p=0.003, power 0.950) (see tables 32-33).

Regression co-efficients of OSS (pain) for different tendinopathy groups

	Unstandardized Coefficients			95.0% Confidence Interval for B	
	OSS (pain)	B	Sig.	Lower Bound	Upper Bound
Normal tendons (Constant)	9.71	9.711	<0.001	8.954	10.468
Abnormal/Partial tears	9.94	.224	.679	-.841	1.289
Full-thickness tears (≤2.5cm)	9.13	-.578	.453	-2.092	.937
Full-thickness tears (>2.5cm)	8.59	-1.119	.163	-2.695	.458

Table 32: There was no significant difference in OSS pain scores between different tendon groups

Regression co-efficients of OSS (function) for different tendinopathy groups

	Unstandardized Coefficients			95.0% Confidence Interval for B	
	OSS Function	B	Sig.	Lower Bound	Upper Bound
Normal tendons (Constant)	29.2	1.035	<0.001	.849	1.221
Abnormal/Partial tears	28.7	.161	.240	-.108	.430
Full-thickness tears (≤2.5cm)	27.4	.483	.012	.107	.859
Full-thickness tears (>2.5cm)	26.3	.702	.001	.287	1.117

Table 33: Only FTTs had a significantly lower OSS function score than normal tendons

3: SUMMARY OF RESULTS

Symptom likelihood

There was a statistically significant difference in the likelihood of having symptoms dependent upon the rotator cuff tendon classification. This difference was not dependent upon age, BMI or whether it was the dominant or non-dominant arm that was affected. The relative risks of having symptoms compared to those with normal tendons was 1.97 for abnormal or partial tears, 2.20 for full-thickness tears 0-2.5cm, and 4.72 for full-thickness tears >2.5cm ($p<0.001$ in all cases). When using a 3 point change in the OSS to define symptoms the results remained statistically the same though with slightly lower

relative risks (1.79, 2.10, and 3.92 for abnormal tendons/partial tears, full-thickness tears 0-2.5cm and full-thickness tears >2.5cm respectively compared to normal tendons).

Looking at pain in isolation the relative risks were 2.06, 2.35 and 5.50 ($p<0.001$ all cases) respectively. For function the relative risks were 1.87, 2.32, and 4.40 ($p<0.001$ all cases) respectively. These again were not affected by age, BMI or whether it was the dominant or non-dominant arm. Using the 3 point change in OSS to define symptoms relative risks for pain were 1.90, 1.98 and 4.52 respectively, and for function were 1.89, 2.44, and 5.11 respectively.

Symptom severity

Of the shoulders that were symptomatic the mean OSS for normal tendons was 42.4, abnormal/partial tears 42.2, full-thickness tears 0-2.5cm 40.0, and full-thickness tears >2.5cm 38.3. There was only a statistical difference in symptom severity between normal tendons and full-thickness tears >2.5cm, which was also clinically significant with a change in OSS of 4 points. Age, BMI or whether it was the dominant or non-dominant arm had no effect on symptom severity. Looking at pain in isolation there was no statistical difference in pain severity between any of the groups. For function the mean OSS function subset scores were 29.2 for normal tendons, 28.7 for abnormal/partial tears, 27.4 for full-thickness tears 0-2.5cm, and 26.3 for full-thickness tears >2.5cm. The difference was significant for all full-thickness tears, however was only clinically significant for full-thickness tears >2.5cm with a change of 3 points.

4: SHOULDER LATERALITY; THE EFFECT OF THE INDIVIDUAL ON SHOULDER SYMPTOMS

The previous section used all 926 shoulders to demonstrate the relative risk of having symptoms in association with rotator cuff pathology compared to normal tendons, which is consistent with how data has previously been presented in the literature. Using all shoulders in a single-level statistical analysis was possible due to the lack of association between symptoms and whether the data came from the dominant or non-dominant shoulder. The results and p-values remain unaffected compared to a multi-level model that incorporated each individual having two shoulders, only there is a wider confidence interval. However, considering that this study has demonstrated that only 51.6% of all full-thickness tears were symptomatic, we have by using a uni-level model including all shoulders, not taken into account any effect that the individual may have on symptom presentation. This relationship has not previously been explored in the existing literature.

Table 34 shows the relationship between the individual and presence of a full-thickness tear with any association of symptoms. The table represents the number of individuals rather than shoulders in each group, and looks at those with bilateral or unilateral full-thickness tears and or symptoms. With respect to the unilateral full-thickness tears and symptoms the figure in brackets represents the number of structural changes with corresponding symptoms in the same shoulder.

Distribution of individual shoulder symptoms according to the presence of full-thickness tears

	Bilateral No Symptoms	Unilateral Symptoms	Bilateral Symptoms	Total
Bilateral No FTT	226	71	63	**360**
Unilateral FTT	33	25 (21 corresponding)	24	**82**
Bilateral FTT	10	3	8	**21**
Total	269	99	95	**463**

Table 34: The majority of individuals reported the same symptoms in each shoulder irrespective of pathology

Of all shoulders with full-thickness tears 51.6% are symptomatic, as defined in the previous section, but here 60 out of 103 (58.3%) individuals with at least a unilateral full-thickness tear were symptomatic. When looking at the group of 21 individuals with bilateral full-thickness tears, if the individual had no influence on symptoms, and symptoms were attributable to structural change only, one would expect to see that 6 individuals (26.6% (0.516^2)) would have bilateral symptoms, 10 (49.9% (0.516x0.484x2)) would have unilateral symptoms and 5 (23.4% (0.484^2)) would have bilaterally asymptomatic shoulders. However we see in fact that rather than 49.9% having unilateral symptoms only 14.3% do, and over 85% have either bilateral symptoms of bilateral lack of symptoms. This suggests that the individual has an influence on the association of symptoms of rotator cuff tears.

Table 35 shows the relationship between the individual and presence of any tendon abnormality with any association of symptoms. The table represents the number of individuals rather than shoulders in each group, and looks at those

with bilateral or unilateral tendon abnormalities and or symptoms. With respect to the unilateral abnormalities and symptoms the figure in brackets represents the number of structural changes with corresponding symptoms in the same shoulder.

Distribution of individual shoulder symptoms according to the presence of any tendon abnormality

	Bilateral No Symptoms	Unilateral Symptoms	Bilateral Symptoms	Total
Bilateral No Abnormality	131	28	28	**187**
Unilateral any Abnormality	72	34 (30 corresponding)	28	**134**
Bilateral any Abnormality	66	37	39	**142**
Total	269	99	95	**463**

Table 35: The majority of individuals reported the same symptoms in each shoulder irrespective of pathology

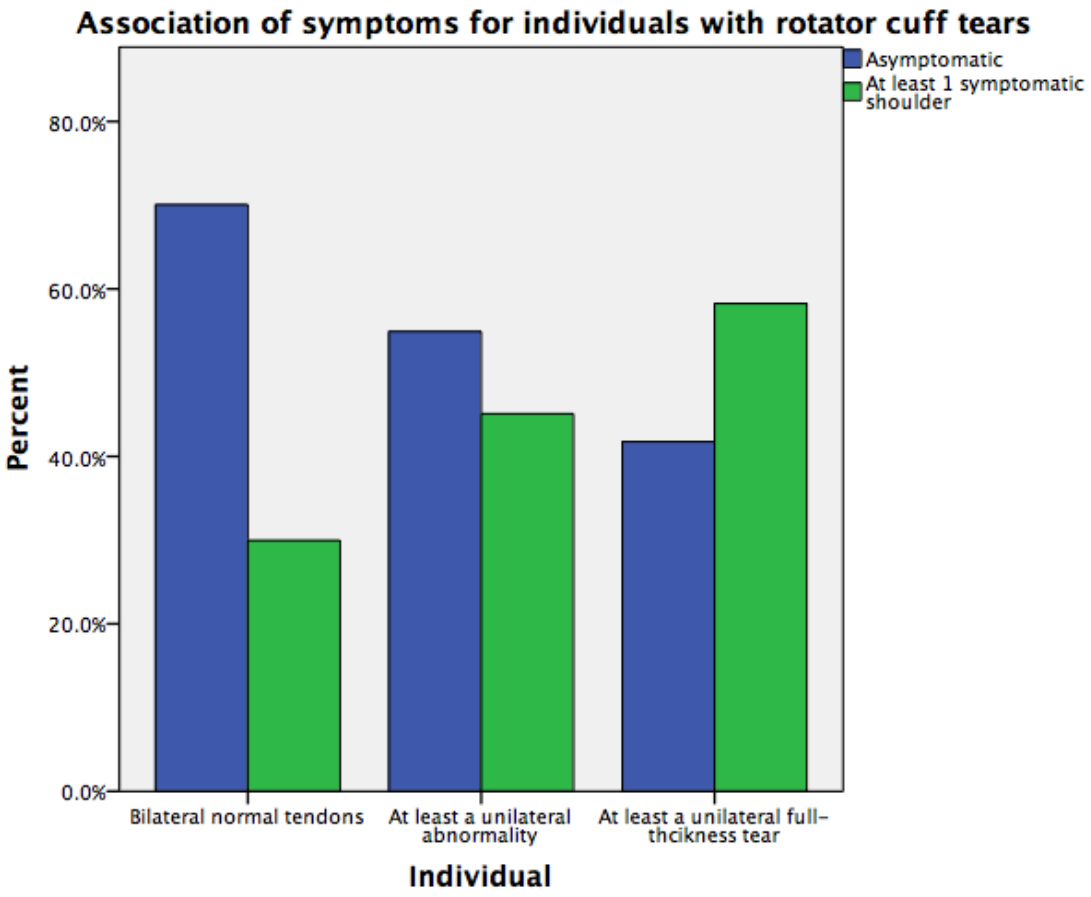
Of all shoulders with tendon abnormalities 41.4% are symptomatic. Therefore when looking at the group of 142 individuals with bilateral tendon abnormalities, if the individual had no influence on symptoms, and symptoms were attributable to structural change only, one would expect to see that 24 individuals (17.1% (0.414^2)) would have bilateral symptoms, 69 (48.6% ($0.414 \times 0.586 \times 2$)) would have unilateral symptoms and 49 (34.3% (0.586^2)) would have bilaterally asymptomatic shoulders. However we see in fact that only 26.1% have unilateral symptoms, and 73.9% have either bilateral symptoms or bilateral lack of symptoms, again suggesting that the individual has an influence on the association of symptoms with rotator cuff pathology.

Of the 463 individuals in the study 103 (22.2%) had at least a uni-lateral full-thickness tear, and 58.3% of these had at least a uni-lateral symptomatic shoulder. 173 (37.4%) had at least a uni-lateral tendon abnormality (but no full-thickness tear), and 45.1% of these had symptoms. 187 individuals had bilaterally normal shoulders and 29.9% of these has current symptoms. This is shown in table 36. The proportion with symptoms was statistically different between groups (Chi^2 test $p<0.001$), as is shown in graph 42.

Proportion of symptomatic shoulders according to the individual and tendon pathology

	N	At least one Symptomatic shoulder	% with symptoms
Bilateral normal tendons	187	56	29.9%
At least one tendon abnormality (no tear)	173	78	45.1%
At least one Full-thickness tear	103	60	58.3%
All	463	194	41.9%

Table 36: The likelihood of symptoms increased with the presence of at least an abnormality or full-thickness tear



Graph 42: The likelihood of symptoms increased with the presence of at least an abnormality or full-thickness tear

Binary logistic regression analysis after adjustment for age, and BMI (No confounding effect p=0.295/0.124) gave the relative risk of having at least one symptomatic shoulder of 1.49 and 1.97 for those with at least a unilateral abnormality or unilateral full-thickness tear compared to those with bilaterally normal shoulders. The model correctly predicted 62.6% of symptom outcomes correctly as shown in table 37.

Regression analysis of having at least one symptomatic shoulder

	OR	RR	Sig.	95% C.I.for OR	
				Lower	Upper
Constant (Bilateral normal tendons)	0.423		<0.001		
Tendon Classification			<0.001		
At least one tendon abnormality (no tear)	1.867	1.483	0.006	1.200	2.904
At least one full-thickness tear	3.352	1.968	<0.001	2.003	5.609

Table 37: The relative risk increased with at least an abnormality and again with at least a FTT

5: SHOULDER PAIN AND PRIMARY CARE USE

To determine the impact that shoulder symptoms in association with rotator cuff pathology has on clinical healthcare services we looked not only at the association of rotator cuff tears with shoulder symptoms, but shoulder symptoms combined with the need to seek medical advise or treatment. The percentages with current or previous shoulder symptoms, along with the percentages of those who sought medical advice for their shoulder symptoms are shown in the following tables. 'Current pain' was defined as per earlier in the chapter, as pain in the last 4 weeks according to the OSS pain scale. Pain, past or present was asked using a binary outcome question 'have you ever had pain or functional problems with your shoulder'?

Tables 38-39 show the percentages of shoulders with pain past or present and the proportion seeking medical advice. Table 38 shows all 926 shoulders in the study divided into abnormalities and full-thickness tears, and table 39 divides them into tendinopathy groups.

Proportion of shoulders with pain, past or present

	Current pain		Pain past or present		Overall % seen GP
	%	% of which seen GP	%	% of which seen GP	
All shoulders (n=926)	25.8	50.0	37.4	52.0	19.4
All tendon abnormalities (n=418)	35.6	51.0	45.5	51.6	23.4
All Full-thickness tears (n=124)	46.0	57.9	55.6	58.0	32.3

Table 38: The proportion with pain increases as tear stage increases, yet the proportion seeking GP advice remains constant for each group

Proportion of shoulders with pain, past or present

	Current symptoms		Symptoms past or present		Overall % seen GP
	%	% of which seen GP	%	% of which seen GP	
Normal tendons (n=508)	17.7	48.9	30.7	52.6	16.1
Abnormal/ partial tears (n=294)	31.3	48.9	41.2	47.6	19.7
FTT 0-2.5cm (n=85)	35.3	46.7	44.7	47.4	21.2
FTT >2.5cm (n=39)	69.2	70.4	79.5	71.0	56.4

Table 39: **The proportion with pain increases as tear stage increases, yet the proportion seeking GP advice remains constant for each group, except for the FTT>2.5cm group which is higher**

Here as we have previously demonstrated the proportion with pain, past or present, increases in accordance with increasing rotator cuff pathology, and also the proportion seeking medical advice also increases (Chi^2 test $p<0.001$). However if we only look at the painful shoulders the proportion seeking medical advice was not statistically different between each tendinopathy group (Chi^2 $p=0.066$), despite it appearing that those with painful large full-thickness tears were more likely to seek medical advice compared to small full-thickness tears. There was also no difference in the likelihood of seeking medical advice in painful shoulders with or without full-thickness tears (Fishers exact test $p=0.116$).

A multivariable regression model using all shoulders was used to predict the likelihood of attending a GP for shoulder pain. Age, BMI, and whether it was the dominant or non-dominant arm did not have any statistical difference

(p=0.787/0.964/0.793). Tendon classification did have an effect that was statistically significant for the large full-thickness tears that had a relative risk of attending the GP for shoulder pain of 3.5 compared to normal tendons (see table 40).

Regression analysis of GP attendance with a painful shoulder

	OR	RR	Sig.	95% C.I.for OR	
				Lower	Upper
Constant (normal tendon)	0.191		<0.001		
Abnormal tendon / Partial tear	1.306	1.245	0.164	0.897	1.901
Full-thickness tear 0-2.5cm	1.405	1.319	0.247	0.789	2.502
Full-thickness tear >2.5cm	6.770	3.510	<0.001	3.401	13.479

Table 40: The likelihood of attending the GP with shoulder pain was only higher for those with large FTTs compared to normal tendons

Table 41 represents the above data by the individual rather than all shoulders, and therefore includes 463 individuals who are grouped by the presence of structural pathology and the presence of pain in either arm.

Proportion of individuals with pain past or present

	Current Pain (either shoulder)		Pain past or present (either shoulder)		Overall % seen GP
	%	% of which seen GP	%	% of which seen GP	Total
All individuals (n=463)	41.9	44.8	55.7	50.8	28.3
Bilaterally normal tendons (n=187)	29.9	41.1	48.1	48.9	23.5
At least one abnormality (no tear) (n=173)	45.1	41.0	57.2	46.5	26.6
At least one full-thickness tear (n=103)	58.3	53.3	67.0	59.4	39.8

Table 41: The proportion with pain increases with tear stage severity, but the proportion seeking medical advice does not differ between groups

This table again demonstrates that the proportion of individuals with painful shoulders increases for those with at least one tendon abnormality (with no tear) and again for those with at least one full-thickness tear. Accordingly the percentage seeking medical advice therefore increases and was significantly different between groups (Chi2 test p=0.005). However when only looking at the cohort who have at least one painful shoulder there was not a statistical difference between groups (Chi2 test p=0.179).

A multivariable regression model using all individuals was used to predict the likelihood of attending a GP for shoulder pain. Age and BMI did not have any statistically significant effect (p=0.816/867). However the presence of at least one full-thickness tear had a relative risk of 1.6 compared to those with normal tendons of attending the GP (see table 42). There was no statistical difference in relative risk of those with any tendon abnormality compared to those with bilaterally normal shoulders.

Regression analysis of GP attendance with a symptomatic shoulder

				95% C.I.for OR	
	OR	RR	Sig.	Lower	Upper
Constant (bilaterally normal tendons)	0.296		<0.001		
Tendon classification			<0.001		
At least one tendon abnormality	1.180	1.123	0.502	0.727	1.915
At least one full-thickness tear	2.179	1.634	0.004	1.282	3.703

Table 42: Overall individuals with a full-thickness tea were 1.6 times more likely to seek medical advice for shoulder pain than those with normal tendons. Those with a tendon abnormality were not more likely to seek advice compared to those with normal tendons.

5 Discussion

This study has shown using a large general population cohort that there is an association between rotator cuff tears, and patient perceived symptoms. The existing literature has demonstrated that the clinical presentation of rotator cuff tears is varied and it is known that a proportion of tears can be asymptomatic. Our study has supported this and has shown that in a general population of elderly women 48.4% of full-thickness rotator cuff tears were asymptomatic. The only other population-based study looking at symptom association with full-thickness tears was that by Yamamoto et al. (76) using a mountain cohort in Japan. They found that 34% of full-thickness tears were symptomatic. However, unlike this study, it was not a general population cohort that has been characterised as representative of western society, and secondly any individual who had restricted shoulder movements or had had any previous treatment to the shoulder was excluded, thus there was a selection bias that would result in a lower association of symptoms. This study was not subject to such selection bias.

This study has demonstrated that as rotator cuff pathology increases the proportion of individuals reporting any form of symptoms increases. The percentages of shoulders with symptoms were 37.1% for abnormal tendons or partial tears, 41.2% for full-thickness tears 0-2.5cm, and 74.4% for full-thickness tears> 2.5cm (51.6% for all full-thickness tears). This was compared with the baseline of 22.8% of shoulders with normal tendons. The relative risks of symptoms compared to normal tendons were 1.97, 2.20 and 4.72 for each group respectively. Statistically these risks were all significantly different relative to

normal tendons, but there was no significant increase in risk between the abnormal tendon group and the small full-thickness tear group. The relative risk of symptoms only significantly increased once tears were greater than 2.5cm.

This agrees with existing studies (62,66,67), which not only show that tears are associated with symptoms but have also suggested that tear size effects the likelihood of the presence of symptoms. These previous studies however, have been subject to substantial selection bias that may have increased the association between rotator cuff tear and symptoms. This in particular is the case with the Washington series investigated by Yamaguchi et al, because the cohort being investigated were actively being treated for a contralateral symptomatic rotator cuff tears. Thus although it was demonstrated that there was an increased tendency to develop symptoms in larger tears, we know that that individual already had a symptomatic cuff tear on the contralateral side. Thus if the individual themselves had an effect on symptom presentation rather than the tear in isolation one would expect that there would be a greater tendency for symptoms the shoulder of that particular individual. This is likely to introduce significant bias as we know that 52% of all tears in this cohort were in fact asymptomatic, including some large tears, so there must be extrinsic additional factors influencing the development of symptoms. This study has addressed this issue and also demonstrated the significance of such selection bias. This study is the first study to look at the relationship between rotator cuff pathology, symptoms and the individual rather than just the shoulder. We have demonstrated that the individual does have an effect on the likelihood of symptom presentation in association with rotator cuff tendinopathy. The

distribution of symptoms in those with full-thickness tears was not random between all shoulders and tended to cluster like symptoms in the shoulders of individuals such that 85% of those with bilateral full-thickness tears either had bilateral symptomatic shoulders or bilateral asymptomatic shoulders. Thus overall looking at individuals we found that of those with at least a unilateral full-thickness tear 58.3% had at least one painful shoulder compared to 29.9% who had bilaterally normal shoulders, or 45.1% with bilateral abnormalities. This created a relative risk of symptoms in the individual of 1.48 and 1.97 respectively for the presence of symptoms in those with tendon abnormalities or full-thickness tears respectively.

This study is the first study to demonstrate the relationship between rotator cuff tears, and indeed stages of rotator cuff pathology, with symptoms in a general population cohort, and is the first study that has explored this relationship in terms of the individual rather than just the shoulder. Furthermore it is strengthened by the lack of selection bias in such a cohort, as the cohort was not initially intended to study shoulder pain.

When comparing symptom severity across different stages of rotator cuff tear, due to the strong negative skew of the OSS, which was our outcome measure it was not possible to compare scores across the whole cohort (see appendix 1, section 3:2 for full details). Thus it was only possible to study the cohort of patients with symptomatic shoulders and compare the scores of those with rotator cuff pathology to those with normal tendons. Thus the cohort used contained 289 shoulders, of which 116 (40%) had normal tendons on

ultrasound. Thus instantly selection bias was introduced into the cohort as these shoulders had pain for another reason, whether it is due to the rotator cuff such as impingement syndrome or from another shoulder pathology such as degenerative or inflammatory arthropathy, or instability. Thus to make the assumption that cuff tendinopathy has more severe symptoms than the normal tendon group is unfounded. By comparing with symptomatic normal shoulders rather than asymptomatic normal shoulders, will therefore have weakened any association introducing a bias towards the null hypothesis as we have not been able to compare to normal asymptomatic shoulders. We did however find that those with full-thickness tears >2.5cm had a mean OSS score of 38.4, a mean 4.12 point reduction compared to those symptomatic normal tendons. For abnormal tendons and full-thickness tears 0-2.5cm, although mean scores were below 48 (42.1 & 40.2 respectively), these were not significantly less than normal symptomatic tendons. Assuming therefore that symptoms in shoulders with an abnormality were due to the abnormality diagnosed on ultrasound, as is inferred already in this chapter, comparing scores between the abnormal tendons, and small and large tears would be reasonable. So although unquestionably underestimating effect on symptom severity, what we have demonstrated is consistent with previous studies of full-thickness tears which have found that severity of symptoms only increases when tears are large enough to effect multiple rotator cuff tendons (77).

The major strengths of this study are that it uses a large general population cohort and is therefore not subject to the selection biases that have affected the previous studies. The cohort was originally investigated with the primary focus

of osteoporosis, and not shoulder symptoms, thus any participation is not driven by shoulder symptoms and not subject to such selection bias. The large general population cohort has also allowed us to explore the relationship between the individual and shoulder symptoms rather than just the shoulder pathology. However there are some potential limitations with the cohort used. Firstly, the cohort can only comment on associations in women aged between 65 and 84. We do not feel that this will have introduced significant selection or detection bias as we found no association between age and symptoms, and previous studies have found no association with either age or sex (67,76). Selection bias could have occurred as only 463/516 individuals who attended the year-20 study underwent a shoulder examination. However individuals were selected at random, and the age and BMI of the groups were not statistically different, therefore this should not have had any effect on results. As the cohort was used in its 20th year it will be subject to survival bias. However, this is unlikely to affect the results here, as there are no known associations between shoulder pain and other medical co-morbidities, and also follow up was not influenced by shoulder symptoms as the main focus of the cohort was osteoporosis.

There are further potential limitations in the methodology of the study. Potential analytic bias is present by having two examiners perform the clinical and ultrasound assessment. This was addressed with a small inter-observer reproducibility study and a comparison of the cohorts examined, which were not significantly different (see appendix 2). Satisfactory intra-observer reproducibility studies and validation studies were performed to demonstrate the minimal effect of intra-observer analytic bias (see appendix 2). There is also

a potential for reporting bias as the same examiner has performed a musculoskeletal examination including the OSS, followed by the ultrasound examination, and was not blinded to the OSS scores prior to performing the ultrasound scan. If as our hypothesis suggested that rotator cuff tears are associated with symptoms, in light of a symptomatic shoulder the examiner may have been more likely to interpret an abnormality on ultrasound, or a normal tendon if the shoulder was asymptomatic, which would strengthen any association between the two. This however, is likely to have effected the marginal ultrasound scans between normal and abnormal more than those between normal and greater pathology, i.e. full-thickness tears, and thus is likely to have less bias on the results presented for the full-thickness tears. To avoid this it would have been preferable to have two independent blinded examiners but this was not possible due to resources available. To demonstrate ultrasound-scanning accuracy a learning curve study was undertaken a priori, which demonstrated both examiners had scanning accuracies comparable to those quoted in the literature (see appendix 2).

The definition of symptoms in previous studies varies widely. Some studies use a binary question, whilst others use a scale and dichotomise it as we have done. The point at which a scale is dichotomised is also variable and may reflect all symptoms or only those considered to be clinically significant. This provides a potential analytic bias as any scale too sensitive will increase our association, and any scale not sensitive enough will reduce association, depending upon background noise and overall test sensitivity. Our decision to use the OSS, and in many parts dichotomise it has been explained and justified in appendix 3.

Although our decision to dichotomise the scale at perfect vs. non-perfect scores has been justified in appendix 3, some authors may consider that this difference is too subtle a change and that a clinically relevant change in symptoms would be better represented by a 3 point change in scale rather than a single point. For comparison the same analysis was carried out using shoulders divided into two groups where asymptomatic shoulders were defined as a score of 46-48, and symptomatic shoulders were defined as scores 0-45. The relative risks of symptoms for each tendinopathy group compared to normal tendons were 1.79/2.10/3.92 when using a 3-point change compared to 1.97/2.20/4.72 when using a single point change. All were clinically and statistically significant as we reported.

Overall this study has demonstrated that not only are stages of rotator cuff tear prevalent, and despite 48.4% of full-thickness tears, and 41.7% of individuals with full-thickness tears being asymptomatic, they are associated with symptoms, in particular full-thickness tears >2.5cm. We have shown that once tears become greater than 2.5cm, a value that corresponds to the average width of a normal supraspinatus tendon, tears are significantly more likely to be symptomatic, and also the symptoms are likely to be more severe, though the change in severity seemed to be more driven by functional symptoms rather than pain. Overall 45.1% of shoulders, and 59.6% of individuals in this population cohort had some degree of degenerate rotator cuff pathology. The prevalence of current symptoms in each group ranged from 17.7% in normal tendons through to 69.2% in full-thickness teas >2.5cm, and 29.9% in individuals with normal tendons to 58.6% in individuals with full-thickness tears.

To understand the overall burden that this has on society, and the impact placed on the health service it is necessary to know the number of instances where medical opinion is sought in relation to shoulder symptoms. Of those with current or past shoulder symptoms approximately 50% will have sought medical advice regarding their shoulder symptoms. This 50% includes those with normal tendons through to full-thickness tears 0-2.5cm. The figure then rises to 70% for those with full-thickness tears >2.5cm, which is perhaps reflective that this group is on average likely to experience more severe symptoms. In absolute figures we see that 16.1% of shoulders with normal tendons will seek medical advice, 19.2% of shoulders with abnormal or partially torn tendons, 21.2% of shoulders with full-thickness tears 0-2.5cm, and 56.4% of shoulders with full-thickness tears >2.5cm, equating to a relative risk of having symptoms of 3.17 compared to normal shoulders. In terms of the individual 23.5% of those with bilaterally normal shoulders will seek medical advice for symptoms, 26.6% of those with at least a uni-lateral tendon abnormality, and 39.8% of those with at least a uni-lateral full-thickness tear. The overall impact of this is large, and equates to a total 28.3% of individuals in this this general population cohort having at some point sought medical advice for shoulder symptoms, of which 35.1% had abnormal tendons and 31.3% had full-thickness tears. Overall 8.9% of individuals in the general population have sought medical advice for shoulder symptoms in the presence of a full-thickness tear, and 18.8% of the population for any tendon abnormality.

In conclusion, this general population study has demonstrated that although 48.4% of full-thickness tears are asymptomatic, the overall burden of shoulder pain in association with rotator cuff pathology on individuals, society and health services is significant, and this is particularly so for tears >2.5cm. This important epidemiology greatly contributes to the overall understanding of rotator cuff tears and although as a cohort study cannot imply any causality in the natural history of tears, it suggests that perhaps focuses of management should attempt to prevent tears becoming larger. It has also highlighted significant areas for potential future research. Firstly, novel showing of the effect of the individual on symptom presentation highlights the need to determine risk factors and likely predictors of symptoms. This could include external influences such as smoking and occupation that have previously been looked at. But also potential effects of the individual on symptom presentation that may include the genetics of pain, and also pain processing, whether this be tissue mediated peripherally or centrally mediated by that individual. Secondly as this was a cross sectional study there would be great potential to follow up the cohort longitudinally to look at both tear progression and also symptom progression or development.

3. Motor strength in association with rotator cuff tears

1. Background information

The initial chapters in this thesis have demonstrated that in a normal population of white women aged between 65 and 84, the population prevalence of rotator cuff tears was 22.2%, which significantly increased with age. The population prevalence of a tendon abnormality was 59.5%, again increasing significantly with age. Rotator cuff tears also are likely to be associated with symptoms. The relative risk of having symptoms compared to normal, age matched tendons, was 2.0, 2.2 or 4.7 for abnormal tendons, small full-thickness tears, and large full-thickness tears respectively.

With respect to functional deficits in relation to rotator cuff tears, these can be measured subjectively by the individual, in this case using the functional scale of the Oxford shoulder score, or quantitatively, using myometric strength testing. We have demonstrated that 29.9% of individuals with abnormal tendons report functional deficits, irrespective of pain, rising to 36.5% and 61.5% for those with small and large full-thickness tears respectively. The relative risks for subjective functional deficits are 1.9, 2.3 and 4.4, for abnormal tendons and small to large full-thickness tears respectively. Large full-thickness tears also had a mean 2.9-point deficit on the OSS compared to normal tendons.

From a clinical perspective it is more important to understand the relationship between rotator cuff tears and an individual's subjective opinion of functional deficit and the effects on activities of daily living, however, it is impossible to distinguish functional loss and pain as isolated entities. Although we have split the OSS into separate subscales for any individual with pain we cannot conclude whether the reported functional loss is related to the rotator cuff tear rather than pain. Thus in order to determine the effects of rotator cuff tears on shoulder function and strength we must use a quantifiable measure (myometric strength testing) and look at its relationship with both rotator cuff tears and pain.

Strength testing also remains a key component of many shoulder scores, including the Constant score (34,77-81) and the American Shoulder and Elbow Surgeons Score (ASES) (32). They are used by Orthopaedic surgeons primarily to detect changes in response to treatments, but also by general practitioners as cost effective screening tools to determine those in need of treatment (81,82). On this basis it is important to know the relationship between strength and rotator cuff tear along with age and pain as to how these components will influence a score and thus how it is interpreted.

A number of studies have looked to determine normal scores and shoulder strengths depending on the age and sex of the individual. Strength, and overall scores, have been shown to significantly decrease with age (83) and female gender (81), with a mean decrease in score of 7.5 points in the Constant score for women (exclusively to strength), and a decrease of 3 points per decade of age (due to strength and range of movement).

With respect to the relationship between rotator cuff tears and strength, a number of studies have tried to remove the potential confounding effect of pain by looking only at asymptomatic shoulders. Kim et al. (83) found a reduction in abduction strength in association with large full-thickness tears, but statistical significance could not be demonstrated due to the size of the cohort, which contained only 29 full-thickness tears, of which only 6 were large. Another study from the same group showed no association between rotator cuff tear and external rotation strength (84). One population study has been performed, using a rural Japanese mountain population, to look at the effects of rotator cuff tear on activities of daily living (85). This study assessed function using the Simple Shoulder test (SST) (86), and looked at both symptomatic and asymptomatic shoulders. Strength though was not formally tested; it was defined by asking a single question, as to whether the individual could lift 3.6kg to shoulder height.

No study to date has measured shoulder strength in association with rotator cuff tear in a general population cohort. In doing so we intend to determine whether there is an associated quantifiable loss of strength in association with rotator cuff tear or whether this is a reflection of pain confounding the results. We will also determine how age interacts with the above factors. This information will not only improve the understanding of the clinical impact of rotator cuff tear but also provide important information for interpreting the results of strength testing which is included in many common practice scores.

2. Aims and objectives

1. Test the hypothesis that different stages of rotator cuff tears are associated with quantifiable functional loss, and whether this relationship is confounded by pain.
2. Does quantifiable functional loss correlate with qualitative functional loss as determined using PROMs?

3. Patients and methods

Patients

The study participants were selected from the Chingford Study, a well described prospective population-based longitudinal study of osteoarthritis and osteoporosis, comprising initially of 1003 women, derived from the register of a large general practice in Chingford, North London (see appendix 1). The women aged 44-67 years at baseline are representative of women in the UK general population with respect to weight, height, and smoking characteristics. The study was established in 1989 to study osteoporosis, and at year 20 of the study 446 had shoulder strength data collected and also underwent a shoulder musculoskeletal assessment including the Oxford shoulder score, and bilateral shoulder ultrasound, (Of the original 1003: 158 women had died; 111 were unable to attend; 218 had either moved away, dropped out or were lost to follow up; 52 attended the year 20 visit but did not have a musculoskeletal assessment due to absence of the assessor; 17 did not undergo strength testing due to lack of equipment on two of the study days; 1 shoulder assessment document was not completed; and 1 individual could not complete strength testing due to effects of

a hemiplegia). The local ethics committee approved the study and consent was obtained from each woman (Outer North East London Research Ethics Committee (formerly Barking and Havering and Waltham Forest RECs), LREC (R&WF) 96).

METHODS

For each participant age, height measured in cm (to the nearest 0.1cm) in a standing position with shoes removed, using a wall-mounted stadiometer (Leicester height measure, Seca) and weight in kg (to the nearest 0.1kg) by electronic scales with shoes removed were recorded. A fixed SOPP (standard operating procedure protocol) was followed for each musculoskeletal assessment, which included the OSS followed by strength testing. This was followed by ultrasound assessment of the shoulders (left then right).

Strength assessment

Shoulder abduction strength was measured using the Nottingham Mecmesin Myometer (Mecmesin Ltd. UK). Strength was assessed according to recommendations by the European Society of Shoulder and Elbow Surgery (www.secec.org), and is the same technique as described by Constant et al. (29,79), standardised by Bankes et al. (80) in obtaining the Constant shoulder score. The participant was seated with the hips at 90 degrees flexion and both feet flat on the floor. Abduction strength was tested with the arm at 90 degrees in the plane of the scapula (30 degrees forward in the coronal plane), with the elbow straight and the palm facing the floor. Flexion strength was tested with the arm at 90 degrees in the sagittal plane, again with the elbow straight and the palm facing the floor. The individual maintained resisted elevation for 5 seconds,

and the maximum strength recorded. This was repeated 3 times, and the mean maximum strength was recorded for each arm.

Symptom assessment

Symptoms have been defined using the validated Oxford shoulder score (30,31). This was completed by the participant prior to attending the study and was reviewed with the clinician during the assessment to reduce error. Binary symptoms were defined using a dichotomised OSS score where any non-perfect OSS was classified as symptomatic. Where questions are pain specific, the four pain specific questions of the OSS were used as a sub-scale and any non-perfect score in the pain subset was classified as symptomatic, whilst perfect scores were asymptomatic. For justification and validation of the use of the OSS as the primary outcome measure see section 1:3.

Ultrasound examination

The ultrasound examination was performed using the GE voluson i portable ultrasound machine with a 10-16MHz linear probe. A fixed protocol was derived according to the recommendations of the Nuffield Orthopaedic Centre Musculoskeletal Radiology Department, and results were recorded on a predetermined data entry sheet. Tendons were classified into one of 4 working groups based upon ultrasound measurements:

1. Normal tendons
2. Abnormal tendons and partial thickness tears
3. Single tendon full-thickness tears (0-2.5cm)
4. Multi-tendon full-thickness tears (>2.5cm)

Justification of the development and use of the above tendon classification, along with validation studies is detailed in section 1:2.

STATISTICAL METHODS

Age, BMI, arm dominance, and symptom presence were compared across the 4-tendinopathy groups. Wilcoxon rank sum test, one-way ANOVA, and chi-squared tests were used for non-normal, normal and categorical data respectively.

Abduction and flexion strengths were compared across each tendinopathy group using a one-way ANOVA. A uni-variable linear regression model was performed to determine the effect of rotator cuff pathology on strength. A multi-variable linear regression model followed this with adjustment for potential confounders, age, pain presence, and hand dominance, along with interactions between age and tendinopathy. A detailed analysis of how these confounding and interacting factors were determined, and which factors required adjustment for is demonstrated in appendix 2 (section 3:2).

All statistics were performed using IBM SPSS Statistics version 22 (IBM, Armonk, NY, USA).

4 Results

1. Descriptive statistics

446 individuals (891 shoulders) were included in the study. Age was not normally distributed and the median age was 71 (range 65-84). BMI was normally distributed with a mean of 27.8. Of the 516 individuals attending the year 20 visit, the median age was 71 (range 63-85), and mean BMI was 27.8, there was no statistical difference in age or BMI between groups (p=0.222/p=0.136).

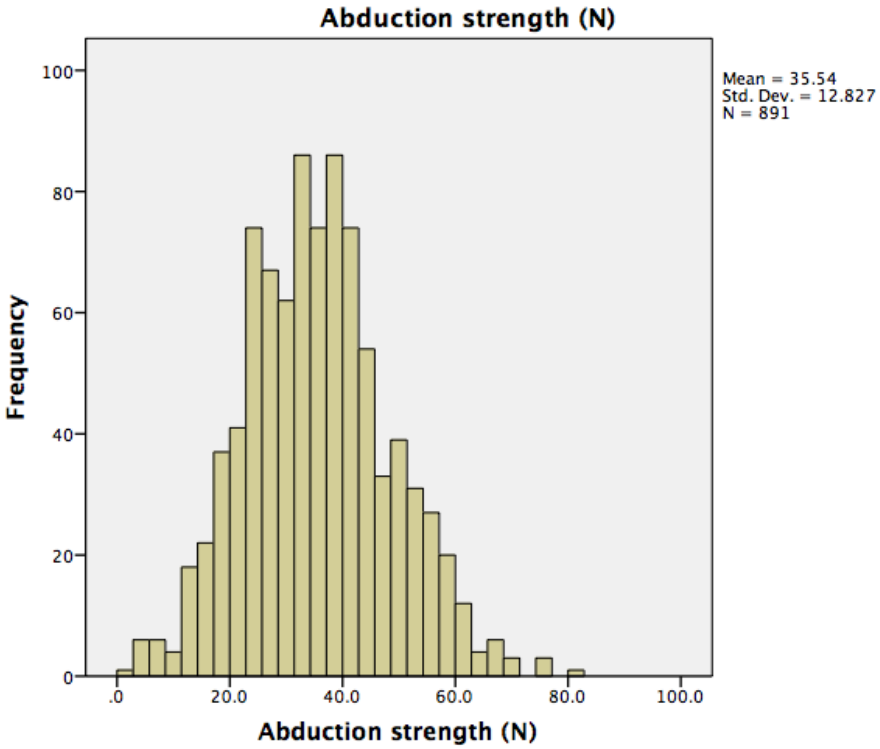
The demographics of the 891 shoulders included are shown in table 43. 489 were diagnosed as normal on ultrasound, 284 abnormal or partial-thickness tears, and 81 full-thickness tears (0-2.5cm), and 37 full-thickness tears (>2.5cm). Age was statistically different between tendinopathy groups (Kruskal Wallis Test p<0.001). There was no statistical difference in BMI between each tendinopathy group (One way ANOVA p=0.078). 402 (90.2%) individuals were right-handed, and 44 (9.8%) were left-handed. There was no difference in right and left-handed individuals in each tendinopathy group (Chi^2 p=0.807), however there was a statistical difference in dominant and non-dominant arms between groups (Chi^2 p=0.047). Of the normal tendons 82 (16.8%) had pain, as defined by the binary OSS pain subset. 89 (31.3%) abnormal/partial tears, 28 (34.6%) full-thickness tears 0-2.5cm, 25 (67.6%) full-thickness tears >2.5cm had pain respectively. This was statistically significant between groups (Chi^2 linear association p<0.001).

Demographics of different tendinopathy groups

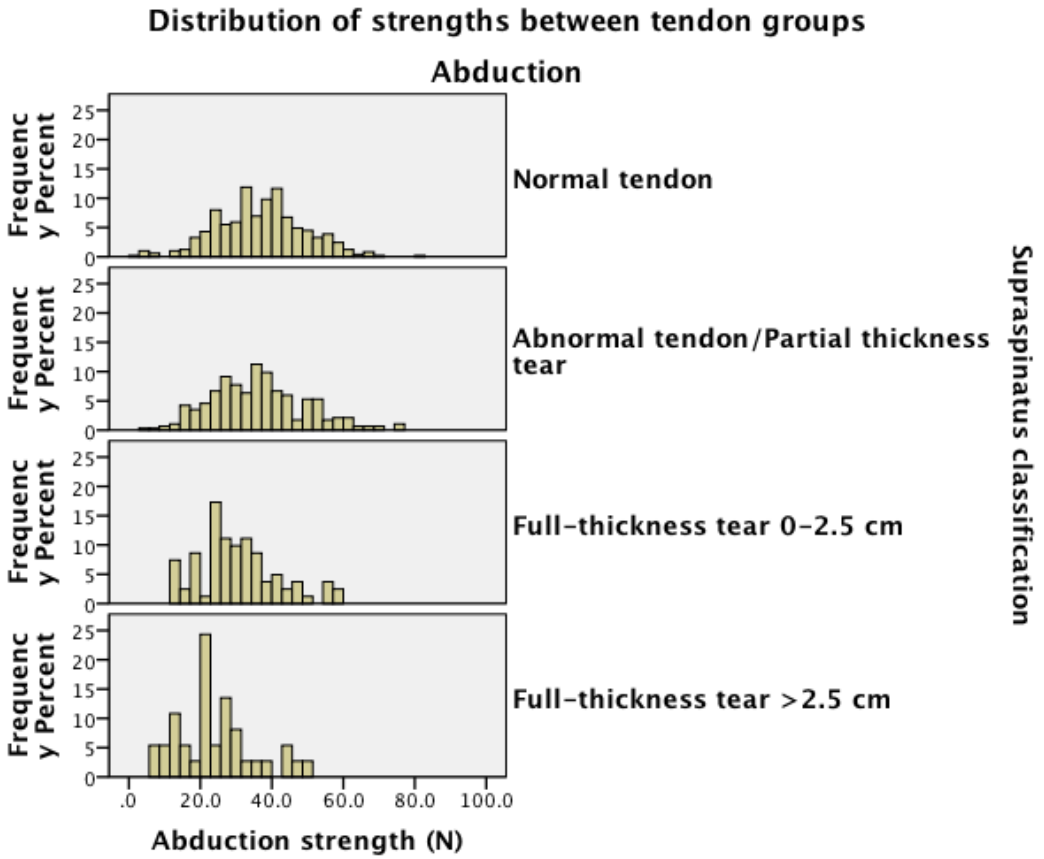
	N	Median age	Mean BMI	% with pain	Dominant arm (%)
Normal	489	70	27.6	16.8%	46.2%
Abnormal/ Partial tear	284	73	28.0	31.3%	52.8%
Full-thickness tears 0-2.5cm	81	74	28.0	34.6%	58.0%
Full-thickness tears >2.5cm	37	74	29.8	67.6%	62.2%
All	891	71	27.8	251%	50.1%

Table 43: The mean age, proportions with pain, and dominant hands increased with tear stage severity. There was no difference in BMI between groups

Shoulder abduction strength was normally distributed (see graph 43). Mean shoulder abduction strength was 35.5N. For normal tendons this was 42.0N, Abnormal/partial tears 39.8N, Full-thickness tears 0-2.5cm 33.4N, and full-thickness tears >2.5cm 24.1N. The difference between groups was clinically significant (One way ANOVA p<0.001) (see graph 44 & table 44).



Graph 43: Graph demonstrating the normal distribution of mean shoulder abduction strength



Graph 44: Mean shoulder abduction strength decreased with increasing tear stage severity

Mean shoulder abduction strength

Tendon quality (n=891)	n	Mean strength (N)	95% confidence intervals	
			Lower	Upper
Normal tendon	489	36.804	35.708	37.901
Abnormal tendon/Partial tear	284	36.311	34.762	37.860
Full-thickness tear 0-2.5 cm	81	30.422	27.929	32.915
Full-thickness tear >2.5 cm	37	24.143	20.486	27.801

Table 44: Mean shoulder abduction strength decreased significantly with increasing tear stage severity

2. Relationship between rotator cuff pathology and quantitative strength

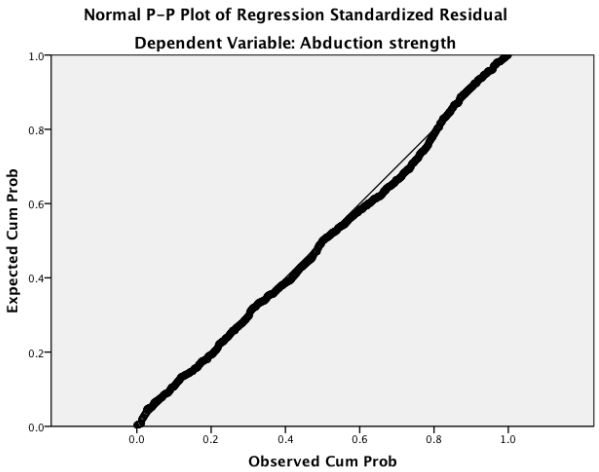
Strength predicted by rotator cuff pathology: Uni-variable model

A uni-variable regression model predicted that there was no reduction in mean shoulder abduction strength between abnormal/partially torn tendons and normal tendons. However full-thickness tears 0-2.5cm and full-thickness tears >2.5cm had mean reductions in shoulder abduction strengths of 6.4N (17.4%) and 12.7N (34.5%) respectively, than the mean score of normal tendons (see table 45). The residual analysis of the model showed that it was working adequately (see graph 45).

Regression co-efficients of mean abduction strength

	Unstandardized Coefficients	Sig.	95.0% Confidence Interval for B	
	B		Lower Bound	Upper Bound
Normal tendons (Constant)	36.804	.000	35.695	37.914
Abnormal and partial tears	-.493	.597	-2.323	1.337
Full-thickness tears 0-2.5cm	-6.382	.000	-9.325	-3.439
Full-thickness tears >2.5cm	-12.661	.000	-16.843	-8.479

Table 45: Full-thickness tears had significantly lower mean abduction strength than normal tendons. There was no weakness associated with abnormal tendons



Graph 45: Residual plot demonstrating that the regression model was performing well with a normal distribution of residuals

Strength predicted by rotator cuff pathology: Multi-variable model

A multi-variable linear regression analysis was performed to show the relationship between shoulder abduction strength and tendinopathy, age, hand dominance and pain (For details of how this model was determined see appendix 2, section 3:2).

The mean dominant arm abduction strength (N) for an individual aged 60-69 with an asymptomatic shoulder and normal ultrasound was 41.4N ($p<0.001$, power 1.0). Pain confounded with a 10.8% (4.5N) mean reduction in strength ($p<0.001$, power 0.96). Non-dominant arms confounded with a 4.9% (2.03N) reduction in strength ($p=0.011$, power 0.72). Overall, the 70-79 age group with normal tendons had a reduction in strength of 10.2% (4.2N) ($p<0.001$, power 0.96), and the 80-89 age group 16.2% ($p<0.001$, power 1.0). Tendinopathy group did not have an isolated effect, but was found to interact with age group. Such, there was no reduction in strength for the 60-69 group irrespective of rotator cuff tear, however for large tears in the 70-79 age group strength was reduced by 39.0% ($p=0.003$, power 0.84), and the 80-89 age group 30.2% for small tears ($p=0.004$, power 0.82), and 32.2% for large tears ($p=0.050$, power 0.49). Details of the regression model and parameter estimates are shown in tables 46-47. Mean shoulder abduction strengths are shown for all shoulders in graph 46. Graphs 47-50 show the effects of symptoms and hand dominance. Tables 48-49 show mean shoulder abduction strengths broken down by hand dominance.

Multi-variable regression model: Variable effects

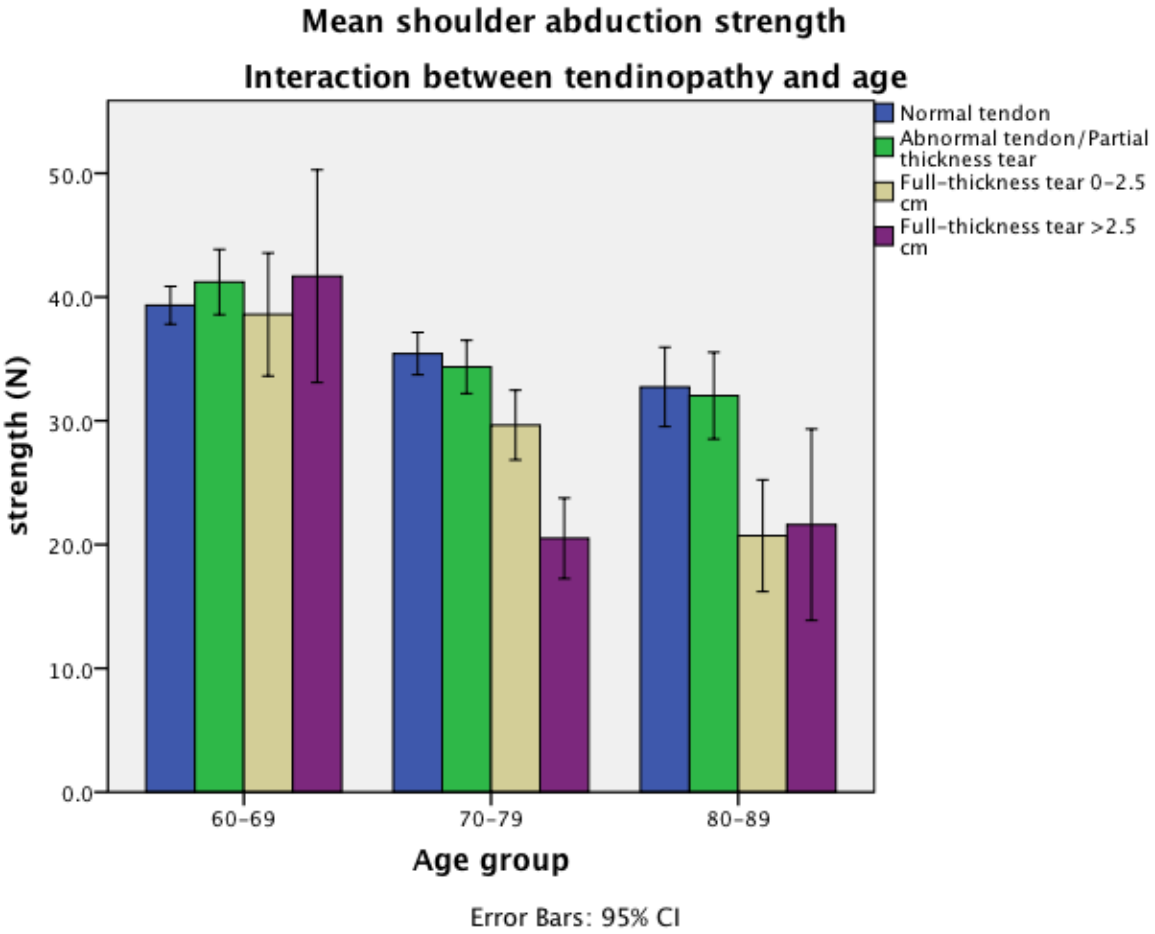
Source	df	Sig.	Observed Power
Corrected Model	13	.000	1.000
Intercept	1	.000	1.000
Tendon classification	3	.000	.974
Age group	2	.000	1.000
Pain status	1	.000	.997
Hand dominance	1	.011	.716
Tendon classification/ age group interaction	6	.009	.891

Table 46: Each input variable or interaction had a significant effect on shoulder abduction strength with >80% power

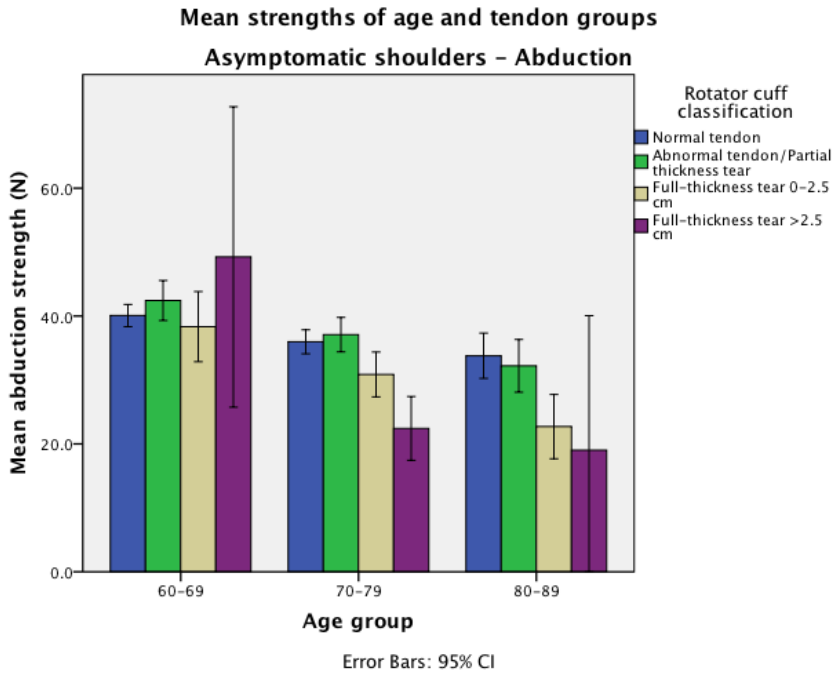
Parameter estimates of multi-variable regression model

	Mean shoulder abduction strength	Sig.	95.0% Confidence Interval for B		Power
			Lower Bound	Upper bound	
Normal tendons 60-69 (Constant)	41.352	.000	39.500	43.203	1.000
Abnormal and partial tears	2.360	.104	-.485	5.205	.369
Full-thickness tears 0-2.5cm	.427	.873	-4.811	5.665	.053
Full-thickness tears >2.5cm	3.695	.454	-5.983	13.373	.116
70-79	-4.222	.000	-6.458	-1.987	.959
80-89	-6.704	.000	-10.284	-3.124	.956
Symptomatic	-4.537	.000	-6.413	-2.662	.997
Non-dominant arm	-2.032	.011	-3.605	-.458	.716
Abnormal and partial tears * 70-79	-2.638	.175	-6.453	1.176	.273
Full-thickness tears 0-2.5cm * 70-79	-5.680	.085	-12.151	.792	.406
Full-thickness tears >2.5cm * 70-79	-16.264	.003	-27.116	-5.412	.836
Abnormal and partial tears * 80-89	-3.107	.254	-8.447	2.232	.207
Full-thickness tears 0-2.5cm * 80-89	-12.499	.004	-21.073	-3.925	.815
Full-thickness tears >2.5cm * 80-89	-13.280	.052	-26.699	.139	.492

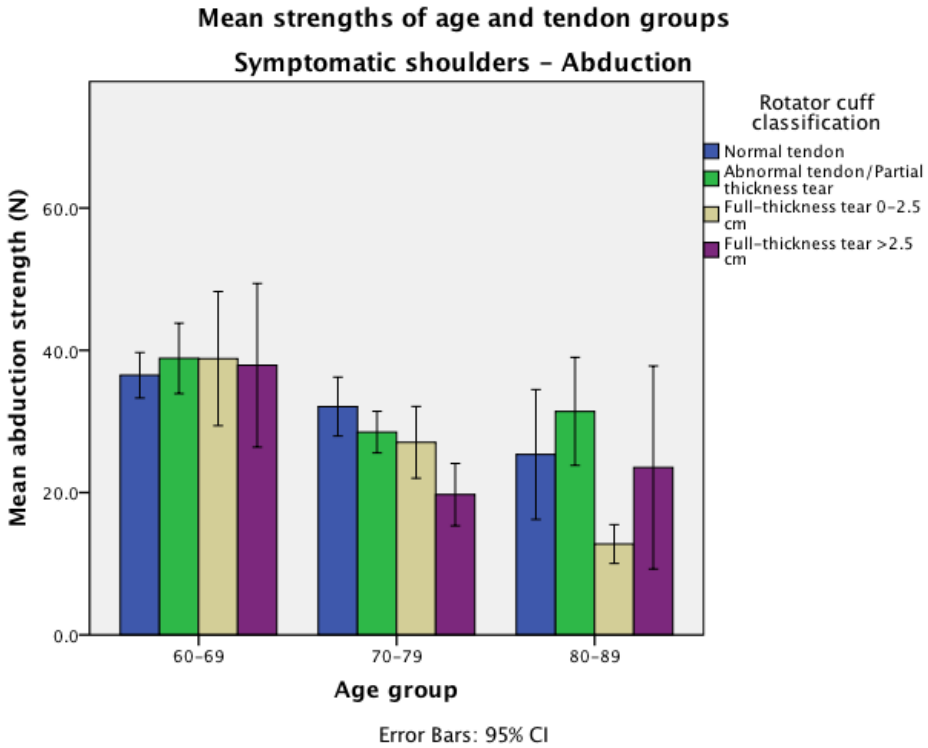
Table 47: Effects of each parameter are shown, of which are summative



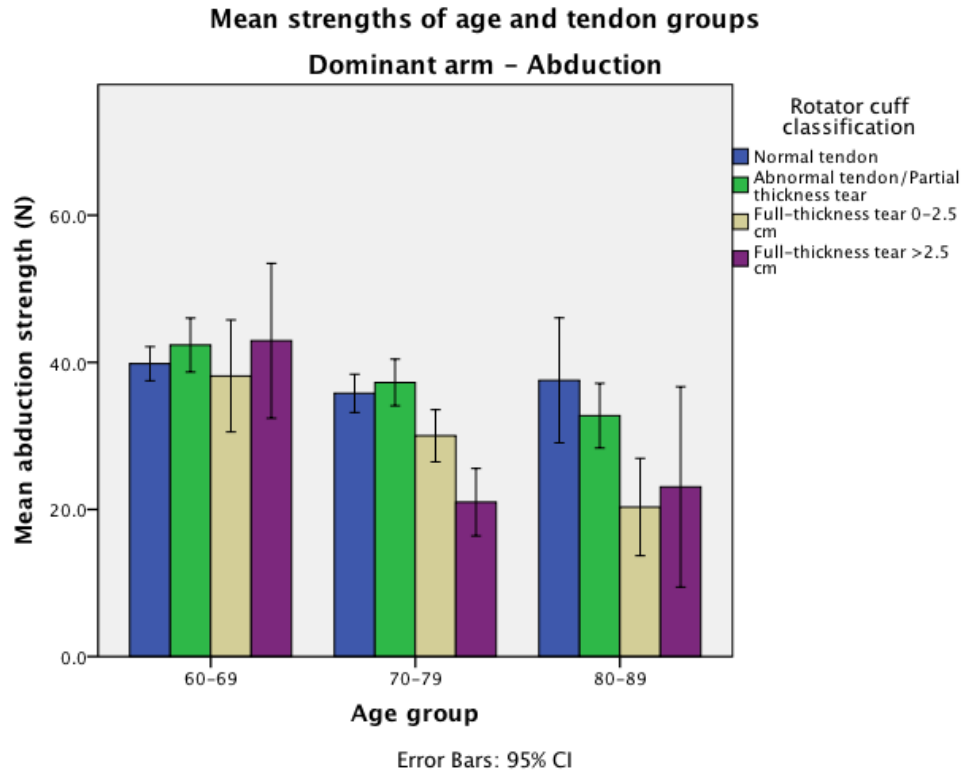
Graph 46: All shoulders included, showing that in the 60-69 age group strength was preserved irrespective of rotator cuff tear, but in the >70's strength decreased in association with rotator cuff tear



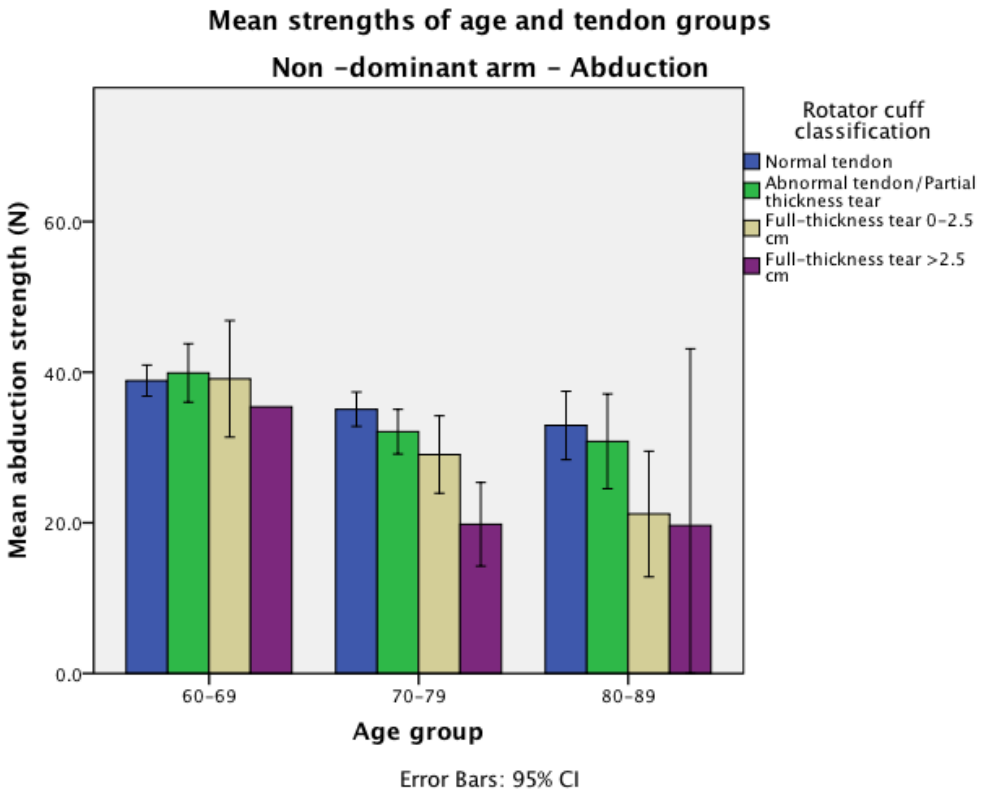
Graph 47: The overall pattern is the same as graph 46, but overall the mean strength in asymptomatic shoulders was higher than symptomatic shoulders, as seen when compared to graph 48.



Graph 48: The overall pattern is the same as graph 46, but overall, the mean strength in asymptomatic shoulders was higher than symptomatic shoulders, as seen when compared to graph 47.



Graph 49: The overall pattern is the same as graph 46, but overall, the mean strength in dominant shoulders was higher than in non-dominant shoulders, as seen when compared to graph 50.



Graph 50: The overall pattern is the same as graph 46, but overall, the mean strength in non-dominant shoulders was lower than in non-dominant shoulders, as seen when compared to graph 49.

Mean shoulder abduction strength according to age and tear stage severity

Dominant arm

		Age groups at time of shoulder exam							
		60-69		70-79		80-89		Total	
		Mean	Count	Mean	Count	Mean	Count	Mean	Count
Normal tendon	Asymptomatic	40.2	76	35.9	89	35.3	18	37.6	183
	Symptomatic	38.5	23	35.4	18	32.1	1	37.0	42
Abnormal tendon / Partial thickness tear	Asymptomatic	42.7	32	39.3	48	33.0	25	38.8	105
	Symptomatic	41.9	21	31.3	16	32.1	8	36.4	45
Full-thickness tear 0-2.5 cm	Asymptomatic	37.5	4	31.1	18	22.6	6	30.2	28
	Symptomatic	38.5	8	27.9	9	13.4	2	30.8	19
Full-thickness tear >2.5 cm	Asymptomatic	49.3	2	21.0	5	28.2	1	28.9	8
	Symptomatic	38.7	3	21.0	9	21.4	3	24.6	15
Total	Asymptomatic	41.0	114	35.9	160	32.5	50	37.1	324
	Symptomatic	39.8	55	30.3	52	27.1	14	34.3	121

Table 48: Effects of confounding and interacting factors determined the mean shoulder abduction strength for each group

Mean shoulder abduction strength according to age and tear stage severity

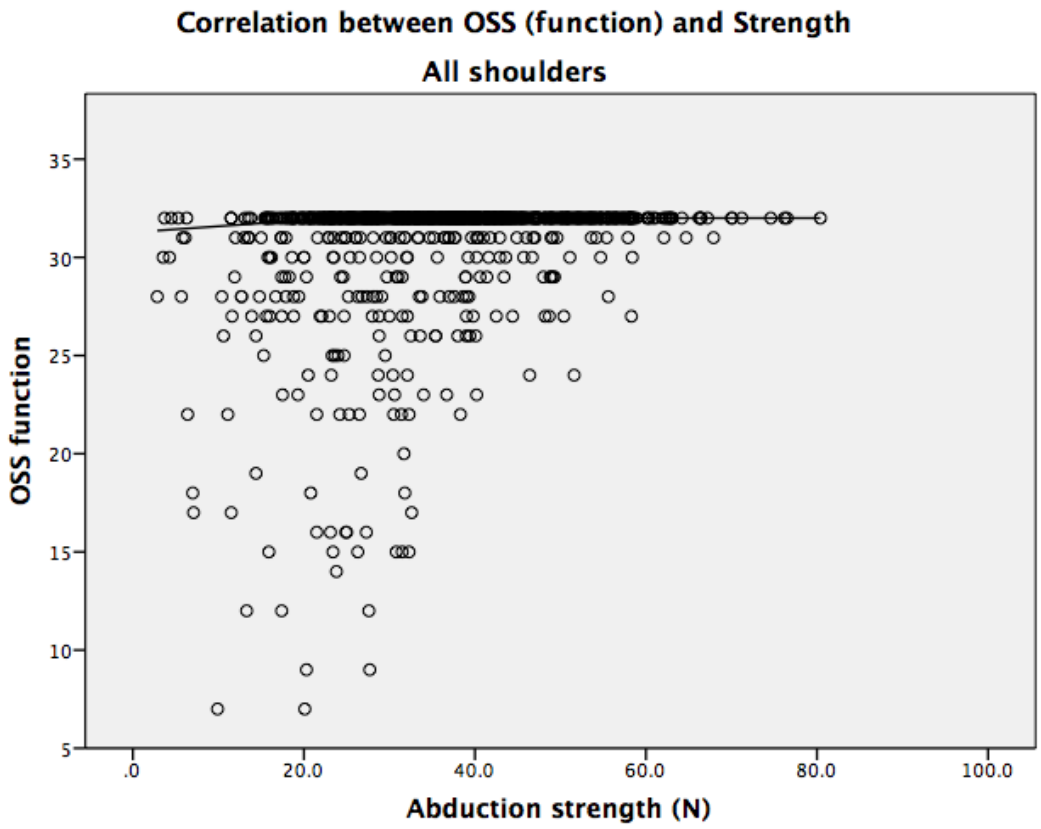
Non-dominant arm

		Age groups at time of shoulder exam							
		60-69		70-79		80-89		Total	
		Mean	Count	Mean	Count	Mean	Count	Mean	Count
Normal tendon	Asymptomatic	39.9	91	36.1	102	32.8	29	37.2	222
	Symptomatic	34.3	21	27.9	14	24.0	5	30.8	40
Abnormal tendon / Partial thickness tear	Asymptomatic	42.2	33	34.6	42	31.0	15	36.8	90
	Symptomatic	34.0	13	26.8	26	30.4	5	29.3	44
Full-thickness tear 0-2.5 cm	Asymptomatic	38.8	7	30.5	12	22.8	6	31.0	25
	Symptomatic	39.7	3	25.5	5	11.5	1	28.7	9
Full-thickness tear >2.5 cm	Asymptomatic	.	0	26.1	2	14.5	2	20.3	4
	Symptomatic	35.4	1	18.3	8	30.0	1	21.1	10
Total	Asymptomatic	40.5	131	35.1	158	30.4	52	36.5	341
	Symptomatic	34.7	38	25.7	53	26.1	12	29.0	103

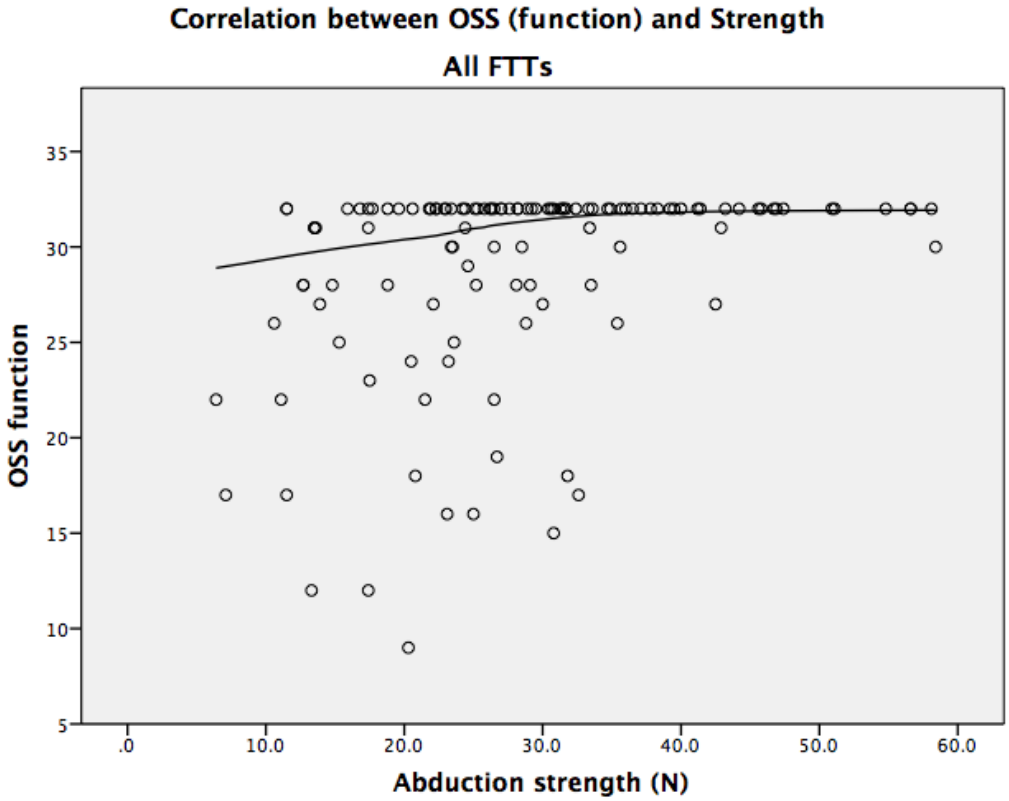
Table 49: Effects of confounding and interacting factors determined the mean shoulder abduction strength for each group

3. Correlation between quantitative strength and patient reported function

Correlation was sought between quantitative strength and the functional component of the Oxford shoulder score. Spearman's correlation coefficient was 0.272 ($p<0.001$) for all shoulders and 0.413 ($p<0.001$) for all full-thickness tears. Although correlation was not high, graphs 51-52 show a clear trend of those with lower OSSs having lower quantitative strengths. Age groups had no effect on the relationship.

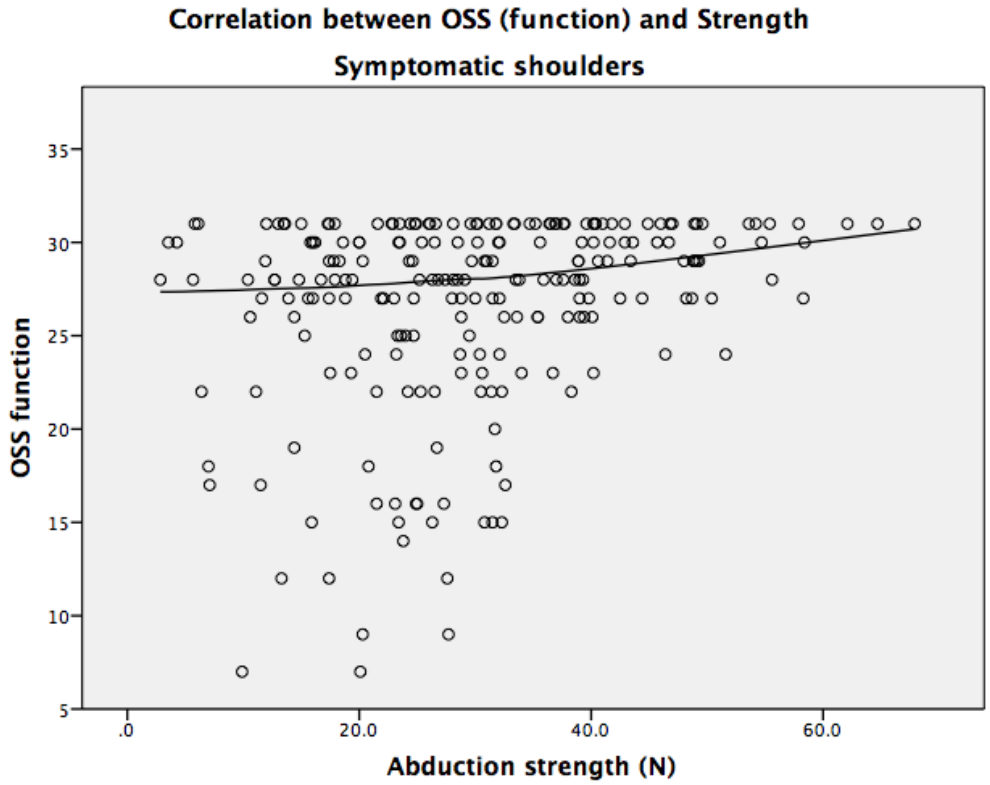


Graph 51: Across all shoulders, those with lower OSSs tended to have lower mean abduction strengths

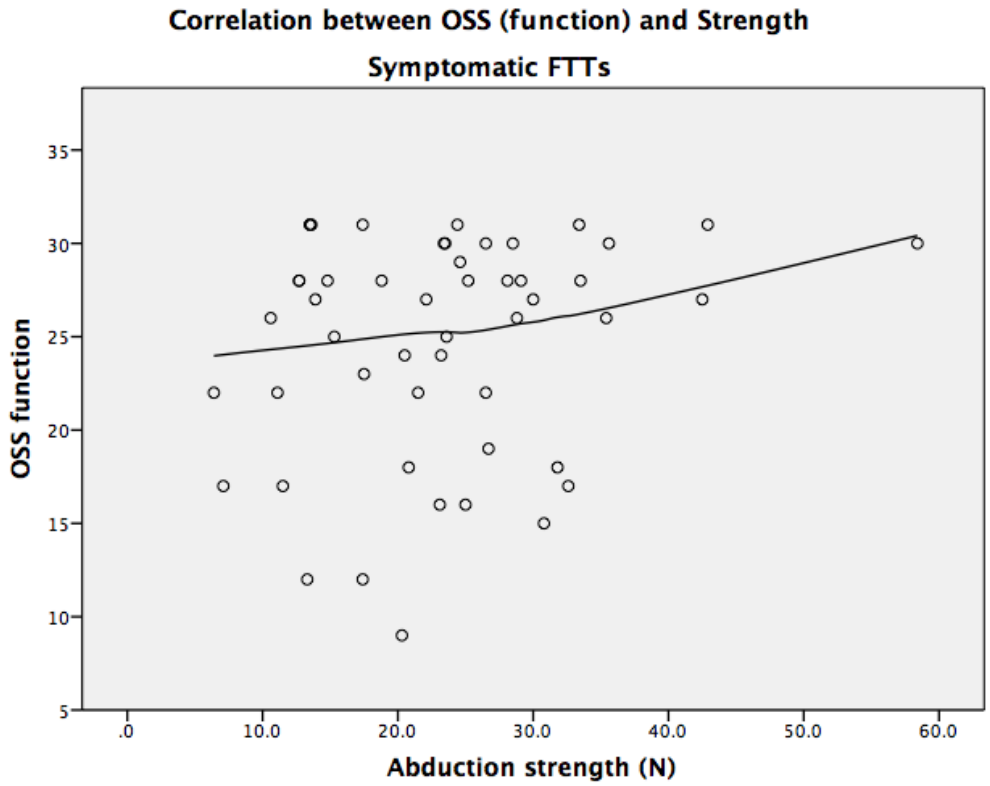


Graph 52: Only looking at FTTs, again those with lower OSSs tended to have lower mean shoulder abduction scores

When looking at only symptomatic shoulders the results were similar. Spearman's correlation coefficient was 0.238 ($p<0.001$) for all shoulders and 0.198 ($p=0.190$) for all full-thickness tears. See graphs 53-54. Age groups had no effect upon the relationship.



Graph 53: Across all symptomatic shoulders, those with lower OSSs tended to have lower mean abduction strengths



Graph 54: Across all symptomatic FTTs, those with lower OSSs tended to have lower mean abduction strengths

5. Discussion

We have shown previously that although not uniformly so, rotator cuff pathology is associated with pain and functional loss, and although an individual's subjective experience is clinically more important in guiding treatments, quantitative assessments are also important. Firstly it is not possible to isolate pain and functional loss, as it may be pain that causes a functional deficit. Secondly as many shoulder scores use quantitative strength assessments as a component, and these can be used to stratify individuals into treatment groups, it is essential to understand how rotator cuff tears and other factors influence results.

This study has shown that rotator cuff tears are associated with loss of shoulder strength, but not as an isolated factor. Using uni-variable models we found the following associations with decreased shoulder motor strength: full-thickness tears 0-2.5cm, and full-thickness tears >2.5cm had a reduction of 17.3% and 23.6% respectively compared to normal tendons ($p<0.001$); age 70-79, and 80-89 had 15.7% and 23.6% reduction respectively compared to those aged 60-69 ($p<0.001$); the reduction due to pain was 13.4% ($p<0.001$); and the non-dominant arm compared to the dominant arm 4.3% ($P<0.001$). However there are interactions between these factors. Rotator cuff tears have a positive association with pain, increasing age, and whether it is the dominant arm, so uni-variable models cannot accurately draw the above conclusions due to bias of an estimator. Multi-variable models to determine the effect of rotator cuff tear on shoulder motor strength showed that overall pain and whether it was the non-

dominant arm confounded the association with reductions of 10.8% ($p<0.001$) and 4.9% ($p<0.001$) respectively. Overall strength in isolation of rotator cuff tear was reduced by 10.2% for ages 70-79, and 16.2% for ages 80-89. The presence of an abnormal tendon or partial thickness-tear had no effect on strength. The presence of rotator cuff tear had no isolated effect on strength, however when interacting with age an effect was present. Irrespective of the presence of any size rotator cuff tear strength was preserved in those aged 60-69, however in those aged 70-89 there was a 33-39% decrease in strength in the presence of rotator cuff tear depending upon the age of the individual and the size of the tear. With adjustments made for the increased prevalence of tears with age, this interaction could be representing a lack of ability to compensate for a tear in the older age groups.

This study also looked at the relationship between quantitative strength and patient reported functional outcomes. Strong correlation would not be expected due to difference in variance of shoulder strengths and OS function scores. Strength in all normal shoulders, and any with pathology, was normally distributed in comparison to the heavy negative skew of OS function scores. The wide variety of strengths from individual to individual would therefore not be expected to correlate with OS function scores. However, there was a clear demonstration that those with significantly lower OS function scores were distributed amongst the lower strength scores. This infers that although reduced strength may be associated in some cases with rotator cuff tears and in some cases with lower functional scores, in isolation it is not a good descriptive tool of functional problems.

The major strength of this study is that it is the first population-based study that has measured shoulder strength quantitatively in a large general population cohort. Previously the only population-based study of shoulder strength has used the simple shoulder test (SST) to define strength, which is a qualitative assessment of strength; this study has assessed strength both qualitatively and quantitatively. It is the only study that has looked at shoulder strength in relation to rotator cuff tears in both painful and pain free shoulders. Thus it is not subject to significant selection or estimator bias so has the capacity to interpret the confounding effects of pain and any potential relationship between the two. This is important information when interpreting data from shoulder scores which include subsets of pain and quantitative function questions. It also overcomes detection bias in that loss of strength is more prevalent in painful shoulders, which in turn are associated with rotator cuff tears, thus excluding painful shoulders reduces the cohort intended to be studied.

Using this particular general population cohort is also a particular strength for this study. This cohort was established with the primary goal of investigating osteoporosis. Therefore even though it was created by invitation from a general practitioners registry it was not subject to shoulder selection bias as its creation. In terms of attrition bias from follow up again it is not selective for those with shoulder problems, which is further strengthened by the lack of known association between osteoporosis and rotator cuff tears. Thus decisions to attend follow up should not be influenced by shoulder problems. The only selection bias for this study is that it includes only elderly women, and thus results are only

applicable to women aged 65-89. As the study is in its 20th year it could be affected by survivor bias. Of the original 1003, 487 did not attend the year 20 visit either due to death, incapacity, or loss to follow up. It is presumed that on average these individuals as a group were older and thus likely to have been weaker, thus adjustment for age in the study was important. They are also likely to have had other medical co-morbidities, which although we do not believe have any association with rotator cuff tears may be associated with generalized loss of strength. Although we therefore may have lost a disproportionate number of weaker individuals this should not have affected the independent variable of rotator cuff tear, but may have caused us to overestimate strength in the older age groups, but should not effect the association with rotator cuff tear

The potential limitations of this study are observer bias and analytic bias. The study was performed by a single observer, and followed a fixed order of data collection which was subjective symptom assessment followed by strength assessment followed by an ultrasound examination. In the presence of a poor strength assessment, particularly if associated with a painful shoulder the observer may have been more likely to diagnose pathology on ultrasound that would strengthen any association. This may also enhance the confounding effect of pain. This could have been reduced using multiple blinded examiners, but this was not possible with the resources available.

Analytic bias may be introduced in the technique of strength measurement. Although we have followed standardised recognised methods of strength assessment it is difficult to ensure that all individuals perform the test in the

designated manor. The method is based upon isolating shoulder strength and relies upon not using major muscles of the torso. However it is possible to shrug the shoulders or to lean away from the myometer when performing the test artificially creating a higher reading. Though this may have happened uniformly across all tendinopathy groups it is possible that those with pain or weakness were more prone to adapting the technique to generate power. If this were the case we would actually have lessened any association between either pain or pathology. In order to reduce this the observer attempted to be as strict as possible when observing techniques used, though with an ageing cohort it was anecdotally thought that comprehension of the instructions was reduced in the older age groups.

In conclusion this study has shown that quantifiable shoulder strength is directly associated with both age and pain with typically a 5% reduction in strength associated with pain and up to a 10% reduction in strength in association with each decile of age beyond 60. We have also shown that there is a relationship between rotator cuff tear and strength, but this is dependent upon age. Thus decreases in excess of one third in strength can be observed in those with rotator cuff tears who are greater than 70 years of age with pain compared to those with normal tendons, no pain and aged 60-69. Overall this has implications on the interpretation of shoulder scores that are being used to assess the rotator cuff and include strength testing, as in many cases reductions in strength are due to factors already measured elsewhere in the test. Thus subjective measures of assessment via PROMs perhaps give a more easily interpretable result.

This study is not only the first to investigate strength, but also demonstrate a novel interaction between rotator cuff tear and age that significantly affects strength. This finding leads to potential future research, and novel therapeutic development. In terms of functional outcomes of rotator cuff tears the role of age related generalised sarcopenia needs to be investigated. Potentially targeting treatments around this could lead to improved functional outcomes in the elderly with large rotator cuff tears.

Concluding remarks and future work

Musculoskeletal pain is the most common source of disability in the western world with a UK budget for management of around £10 billion (87). The shoulder is the third most common site of pain (1), of which rotator cuff tears are the most common source. It is estimated that 20% of the population will encounter shoulder pain at some stage (2), and given the association with age this is going to become more prevalent in an ageing population where it is estimated that by 2035 one quarter of the population will be aged 65 or over (87). Considering the substantial health economic burden the consensus on the management of rotator cuff disease is varied. Despite a lack of evidence, the rates of arthroscopic sub-acromial decompression have risen by 746% in the UK alone over the last 10 years. The rates of rotator cuff repair have risen by over 100% but have fluctuated in both number and according to geographic location(88). This inconsistency in management may well be related to the variation in clinical manifestations of rotator cuff tears and the inherent lack of understanding of the natural history of rotator cuff tears. The overall aim of this study was to gain a broader understanding of the epidemiology of rotator cuff tears in the general population with respect to both the prevalence and the association of symptoms. This will provide a foundation to the understanding of the natural history of rotator cuff tears, which will ultimately have a clinical impact in determining the management and treatment of individuals with shoulder pain, and rotator cuff tears.

The prevalence of rotator cuff tears has been estimated from many studies, though none of these have been general population studies. This study has demonstrated a population prevalence of rotator cuff tears of 22.2% increasing with age from 14.4% for those aged 65-69 to 29% for those age 80-84, with the latter having a 2.3 times relative risk of full-thickness tear compared to the former. The prevalence was not affected by BMI, but there was a 1.64 increased relative risk in the dominant compared to the non-dominant arm. This is in keeping with other data in the literature.

Uniquely this study has demonstrated a previously unidentified group of tendinopathy visible on ultrasound: The footprint abnormality. It became clear that accuracy of diagnosing partial-thickness tears using high definition ultrasound was not adequate, however this group of footprint abnormalities was easily detectable. Moreover the symptom profile associated with these footprint abnormalities was different to normal tendons and full-thickness tears, but the same as partial thickness tears. All of the full-thickness tears diagnosed on ultrasound had this footprint abnormality, which raises the speculation that the footprint abnormality could be a preceding event to a degenerative partial or indeed full-thickness tear. If this were the case then there are implications for potential clinical management and aiding our understanding of the natural history of degenerative rotator cuff tears. Firstly in terms of progression and implications this group would be provide a good longitudinal study to see what happened to the tears over a period of five to ten years, as to whether they progressed and what happened to their symptoms. In the laboratory full-thickness tears have been studied at a cellular level as a result of biopsies around

the time of surgery, and if the same processes were taking place prior to this at the footprint then this could potentially be a novel area to target therapeutic management in order to prevent potential progression towards full-thickness tears. The underlying cellular processes may also help to understand the natural progression and history of the disease. In this study the prevalence of tendon enthesis abnormalities was 59.4%, rising from 51.5% in those aged 65-69 to 72.5% in those aged 80-84, demonstrating the significant proportion of the population affected. Hence this study has highlighted an area of potentially significant future research work.

Aside from prevalence, less well established in the literature is the association of symptoms with rotator cuff tears. This study has demonstrated that shoulders with rotator cuff pathology are associated with an increased likelihood of having symptoms compared to shoulders with normal tendons. We report that overall 51.6% of shoulders with full-thickness tears were symptomatic, compared to 22.8% of normal tendons, rising to 74.4% in shoulders with tears greater than 2.5cm. The relative risk of a shoulder with pathology being symptomatic compared to a shoulder with a normal tendon was 1.97, 2.20 or 4.72 for abnormal tendons, full-thickness tears 0-2.5cm, and full-thickness tears >2.5cm respectively. Other studies have also demonstrated that larger tears are more likely to be symptomatic than smaller tears, but the reason for this has never been identified. This study showed that the clear change in likelihood of symptoms was found at the 2.5cm size, which corresponds to the average width of the supraspinatus tendon. This would suggest that potentially a cause for this was altered shoulder kinematics with multi-tendon tears. Although some studies

have looked at the proximal migration of the humeral head as a predictor of pain, but not found it to be the case, this certainly is an area where future research could be directed. After all in this study 52% of all full-thickness tears remained asymptomatic, which demonstrates that there are other factors involved in patients reporting pain as opposed to simply the pathology. Furthermore, in terms of patient reported severity of symptoms, only tears greater than 2.5cm demonstrated significantly lower OSSs, with a reduction in score of only 4 points, of which the main drive came from reported functional deficit rather than pain, which was no different. Although this area was difficult to study due to statistical modelling of the data we did find that in the general population the severity of pain was not related to the extent of structural damage. Intact but abnormal tendons had the same pain levels as full-thickness tears, and this was a novel finding.

The data presented above is consistent with other studies in the literature. However it does not give the complete picture between an individual, shoulder pathology and the likelihood of symptoms. After all 48.4% of full-thickness tears are asymptomatic. This is the first study that has taken into consideration the individual as well as the shoulder pathology. Accepting that approximately 50% of full-thickness tears are asymptomatic, it would be anticipated that in individuals with bilateral full-thickness tears, 25% would have no symptoms, 25% would have bilateral symptoms and the remaining 50% would have a unilateral painful shoulder. However, this is not the case, and in fact 85% either have bilateral symptomatic or asymptomatic shoulders, demonstrating that the individual has an interaction with the pathology in the presentation of

symptoms. This study has shown therefore that the relative risk of having symptoms in the presence of at least a unilateral full-thickness tear was 1.97 compared to an individual with bilaterally normal shoulders. For those with at least a unilateral tendon abnormality the relative risk was 1.48 compared to normal. 41.7% of individuals with at least a unilateral full-thickness tear had no shoulder symptoms. This data is a novel finding, and demonstrates that there are other factors influencing symptom presentation rather than merely the absolute pathology, particularly at the level of the individual. Over the last decade much work has gone on into furthering the knowledge of pain, and its mediation both peripherally and centrally. In the year 20 visit of the Chingford study quantitative sensory testing of pain, along with data from the PainDetect questionnaire was obtained. Although a preliminary analysis of this was run in association with rotator cuff tears, and did not reveal any significant findings, it clearly is an area to target in the future. Predicting which individuals will get pain may well be pivotal in providing management, after all the apparent same pathology is producing debilitating symptoms for some but normal pain free states for others.

Although using the OSS as the primary outcome measure to reflect the clinical significance of symptoms, in terms of health economics what is relevant is the proportion of individuals who are seeking medical advice for rotator cuff related shoulder pain. This study is the first to have looked at GP attendances with respect to shoulder pain and rotator cuff pathology. In this cohort 28.3% of individuals (19.4% of shoulders) had seen their GP for shoulder pain. In terms of shoulders 16.1% of shoulders with normal tendons had been consulted upon.

This rose to 19.7% for abnormal tendons, 21.2% for full-thickness tears 0-2.5cm, and 56.4% for full-thickness tears >2.5cm. However this in part reflects the increase in likelihood of symptoms with increasing pathology, and in fact when looking only at symptomatic shoulders only around 50% of each group sought medical opinion apart from the larger full-thickness tear group where this rose to just above 70%. In terms of individuals rather than shoulders, 23.5% of those with bilaterally normal shoulders sought medical advice. This rose to 39.8% in those with at least a unilateral full-thickness tear with a relative risk of GP attendance 1.63 times of those with bilaterally normal shoulders. Overall 8.9% of this general population cohort had sought medical advice and had at least one uni-lateral full-thickness tear, and 18.8% with at least a unilateral abnormality, demonstrating the magnitude of the potential baseline health economic burden.

The primary outcome of this study was to look at patient reported symptoms, thus the Oxford shoulder score was used as the primary outcome measure. However many shoulder scoring systems have quantitative strength as an integral part of the score. Although such scoring systems are useful in detecting change they are increasingly being used as screening tools to determine symptom severity. This is the first study to have measured shoulder strength in association with rotator cuff tears in a general population cohort. Multi-variable modelling showed that rotator cuff tears in isolation were not associated with a quantifiable decrease in strength. The direct contributors to loss in shoulder strength were pain (5%), arm dominance (11%), and age (70-79 10%, 80-84 16% decrease compared to those aged 65-69). Full-thickness tears however, were associated with a decrease in strength in those aged over 70. Strength was

preserved in those under 70 irrespective of pathology, but strength was reduced by between 33 and 39% in those aged over 70 in the presence of a rotator cuff tear. This was a major novel finding and suggests that shoulder function and strength is multifactorial. It raises the question as to why the younger patients with the same pathology do not have any shoulder weakness, but older patients do. Is it suggesting that somehow the younger individuals have a mechanism to compensate for a full-thickness rotator cuff tear where as the older individuals cannot? If this was to be the case then this could have potential clinical implications as to how to treat individuals with cuff tears particularly in the older age groups. Are these people suffering from generalised sarcopenia and therefore cannot compensate, or could indeed different measures of treatment be used in isolation of as augmentation to current treatments to improve shoulder strength and function. It would be interesting to look at the rates of sarcopenia in this cohort to see if this has any correlation with our results. If this were the case then certainly areas of future research into preventing this could help in the outcome of rotator cuff tear, or indeed in the rate of presentation potentially. Quantitative strength however was not strongly correlated with functional strength scores, and in isolation is not a good descriptive tool of functional deficit in relation to rotator cuff tears, but it has highlighted some interesting novel findings.

The major strength of this study has been the use of a general population cohort, which has never been used to study the epidemiology of rotator cuff tears to date. This particular cohort has been demonstrated to be representative of the general population in the UK for women aged between 65 and 84 with respect to

weight, height and smoking status. A particular strength of this cohort as well as being a general population cohort is that it was not set up to study shoulder pain or pathology, and was originally set up to study osteoporosis. This means that it is not subject to selection and surveillance bias that other cohorts have been subject to, and will have falsely elevated any prevalence or symptom associations. The potential limitation of the cohort lies in its generalizability. This is a cohort that includes only women aged now between 65 and 84. However as a study of degenerative rotator cuff disease it is an appropriate cohort to use, and is reflective of almost one quarter of the female population in terms of age. The other potential limitation of the cohort is that it was founded 20 years ago and has been subject to loss to follow up through death, increasing morbidity, consent with-drawl and unknown loss. Further potential limitations in the methodology have been recognised and have been addressed extensively within the study and have been discussed previously.

Summary

Overall this study has met the aims and objectives set out at the onset, and identified a number of novel findings, which are listed in below:

1. **Shoulder strength**
    - Rotator cuff tears of all sizes in those aged under 70 were not associated with a loss of shoulder strength.
    - In those aged over 70, strength was reduced by 30% with small and 40% with large full thickness tears.

- Loss in strength was associated a loss of ability to perform activities of daily living but only for large tears.

2. **Symptom severity**
    - Severity of pain was not related to the extent of structural damage (intact but abnormal tendons had the same pain level as large tears).
3. **Impact on health services**
    - 29% had at some time seen their GP in relation to shoulder pain.
        - 10% had seen their GP with pain and a normal rotator cuff on ultrasound.
        - 10% had seen their GP with pain and an abnormal but intact rotator cuff.
        - 9% had seen their GP with pain and a full-thickness rotator cuff tear.

Aside from these we have demonstrated an overall population prevalence of rotator cuff tears in those aged 65-84 of 22% increasing with age. We have further defined the previous un-speculated prevalence of 59.4% of degenerative tendon abnormalities that we propose may precede full-thickness tears in degenerative tendons. We have explored the relationship between pathology and symptoms in both degenerate and torn tendons and have shown a positive association between rotator cuff pathology and symptoms, but have shown that despite this 48.4% of full-thickness tears are asymptomatic. Although the likelihood of developing symptoms is related to the degree of pathology, the severity of symptoms was not. Uniquely this study has recognised the role of the individual in symptom presentation rather than merely the pathology, and still,

we see that only 41.7% of those with a full-thickness tear will have shoulder symptoms. These figures highlight the need for further research, to determine what factors may contribute to individuals being symptomatic. This could potentially provide alternative targets to treatments. In terms of the burden on medical services we have further demonstrated that only between 50% and 70% of those with symptoms will seek medical advice. This corresponds to 8.8% or 18.8% of the population seeking advice for shoulder symptoms in the presence of either a full-thickness tear or tendon abnormality respectively. Lastly the study has demonstrated that not only are rotator cuff tears associated with shoulder symptoms in some individuals but also a loss in quantifiable strength but this is only in those aged over 70, perhaps indicating that in this age group there is no longer the means to compensate for a tear which there is in the under 70's where strength is preserved. Again this highlights an important potential research question. Why are the over 70's seemingly unable to compensate for rotator cuff tears, and could treatments targeting strength be of benefit in this age group?

This is the first study that has been able to demonstrate the epidemiology of degenerative rotator cuff tears in a general population cohort. Although this is a cross-sectional study and cannot comment upon causality it provides the knowledge foundation that is required to help in the understanding of the natural history of rotator cuff tears which ultimately will help in the management and treatment of individuals with degenerative rotator cuff disease.

It further has highlighted some important areas where further research is required. Firstly the importance of the individual in the presentation of symptoms infers that other factors in symptom presentation need to be explored. Many factors such as occupation, sport, and smoking have in part been looked at but the role of the genetics of pain and individual pain processing may be important. This could be tissue mediated at the periphery or more centrally mediated. Secondly, to improve the understanding of the natural history of rotator cuff tears it would be ideal to follow up such a cohort longitudinally to look at both tear progression and symptom development or progression, particularly in the footprint abnormality group. Lastly in terms of strength this study has demonstrated the multifactorial causes of strength reduction and the functional effects to the individual. The significant interaction between age and rotator cuff tear effecting strength seems to infer that age related generalised sarcopenia may have a role in the functional outcome of rotator cuff tears and certainly needs to be researched as this may open alternative treatment pathways.

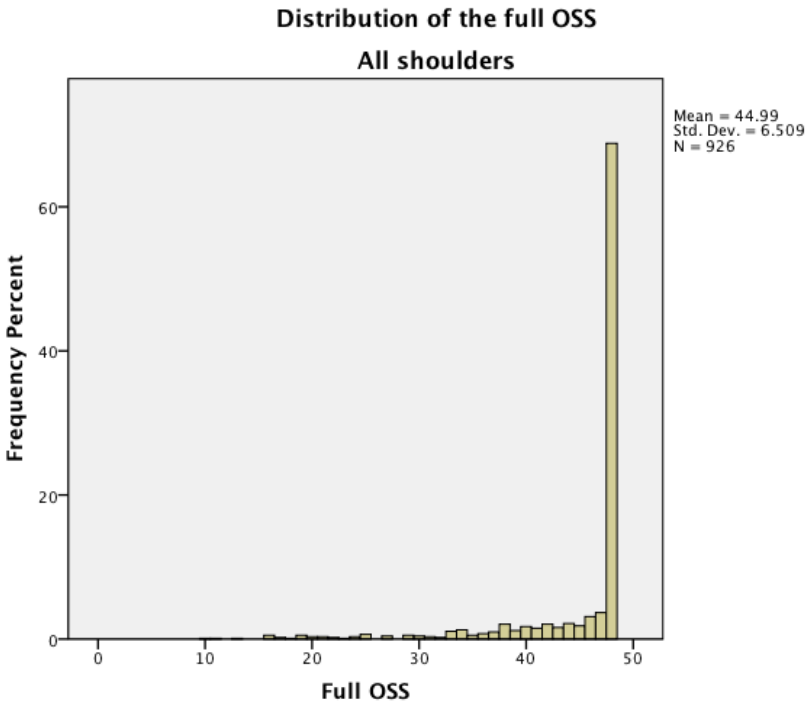
Section 3: Appendices

1: Transformation and grouping of data

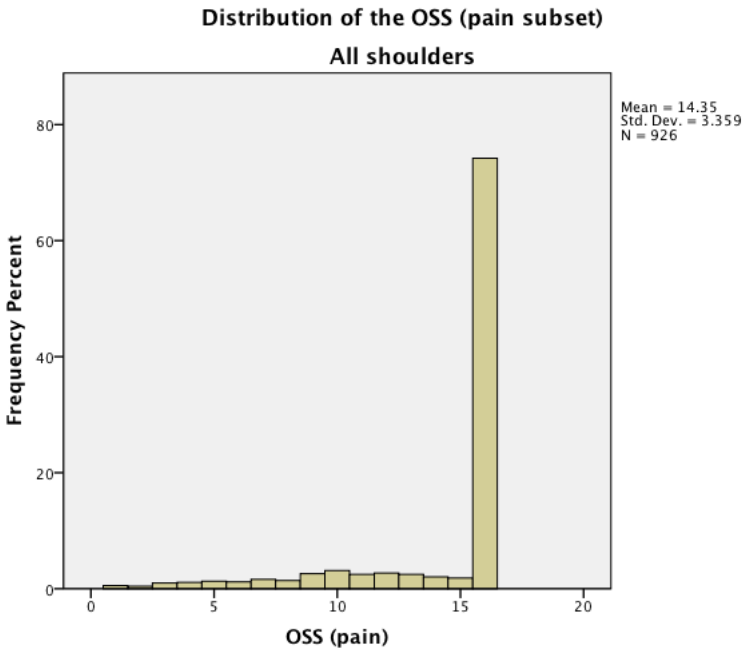
1 Problems with OSS data distribution

The overall aim of the thesis is to determine whether firstly there is an increase in prevalence of symptoms with an increase in rotator cuff pathology, and secondly to determine whether greater pathology is associated with a greater severity of symptoms. The first aim has been answered using the dichotomised OSS and OSS subscales and is justified in section 1.3. However to determine the second aim a scale variable is required. The OSS and OSS subscales were selected for this purpose as they offered the greatest content and construct validity when applied to this study (see section 1.3).

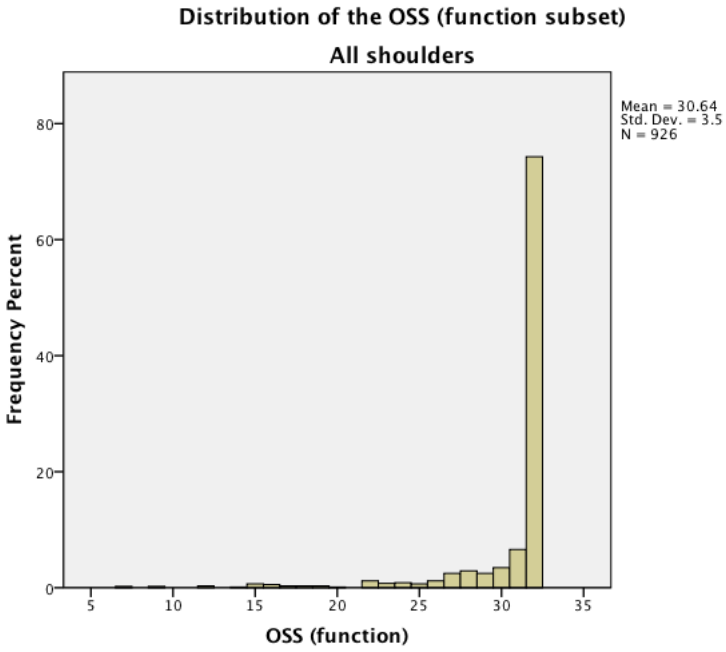
Both the OSS and its subscales however pose problems in answering this question. The disproportionate number of perfect scores means the data is heavily negatively skewed and cannot be normalised with any standard transformations (see graphs 55-57).



Graph 55: Disproportionate number of normal OSS scores gives heavy negative skew



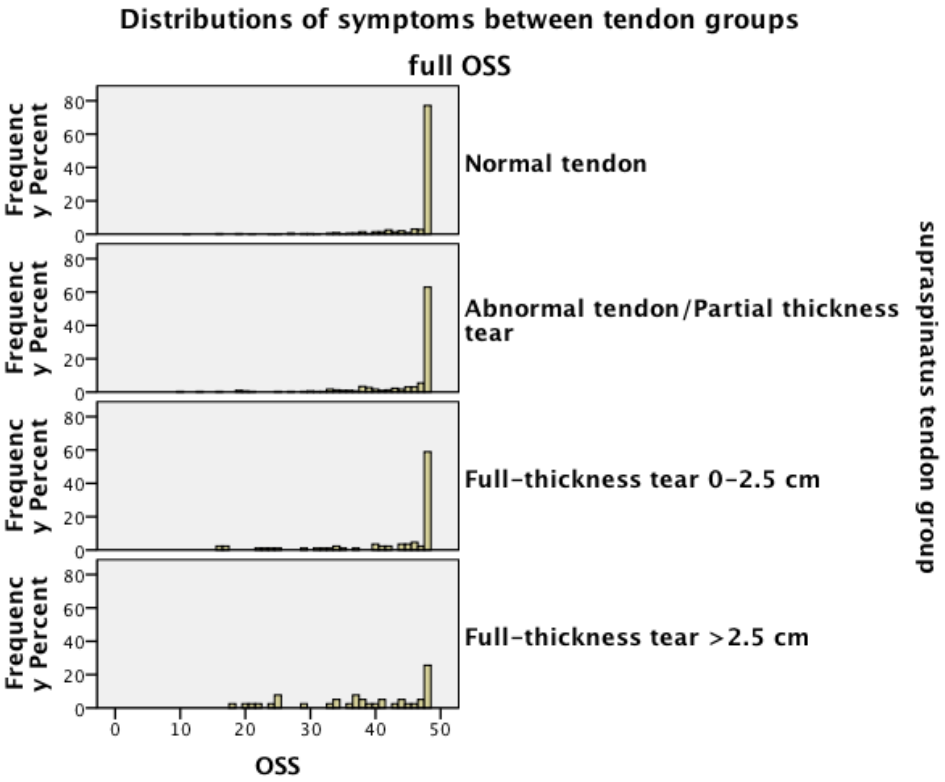
Graph 56: Disproportionate number of normal OSS pain scores gives heavy negative skew



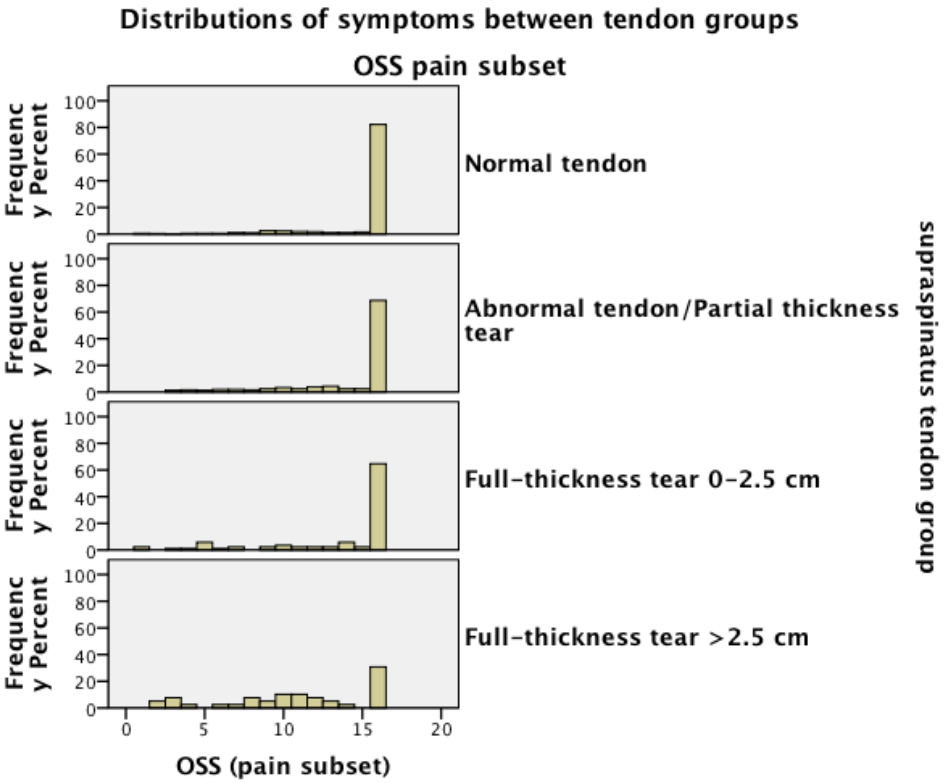
Graph 57: Disproportionate number of normal OSS function scores gives heavy negative skew

COMPARING DISTRIBUTIONS ACROSS DIFFERENT TENDON GROUPS

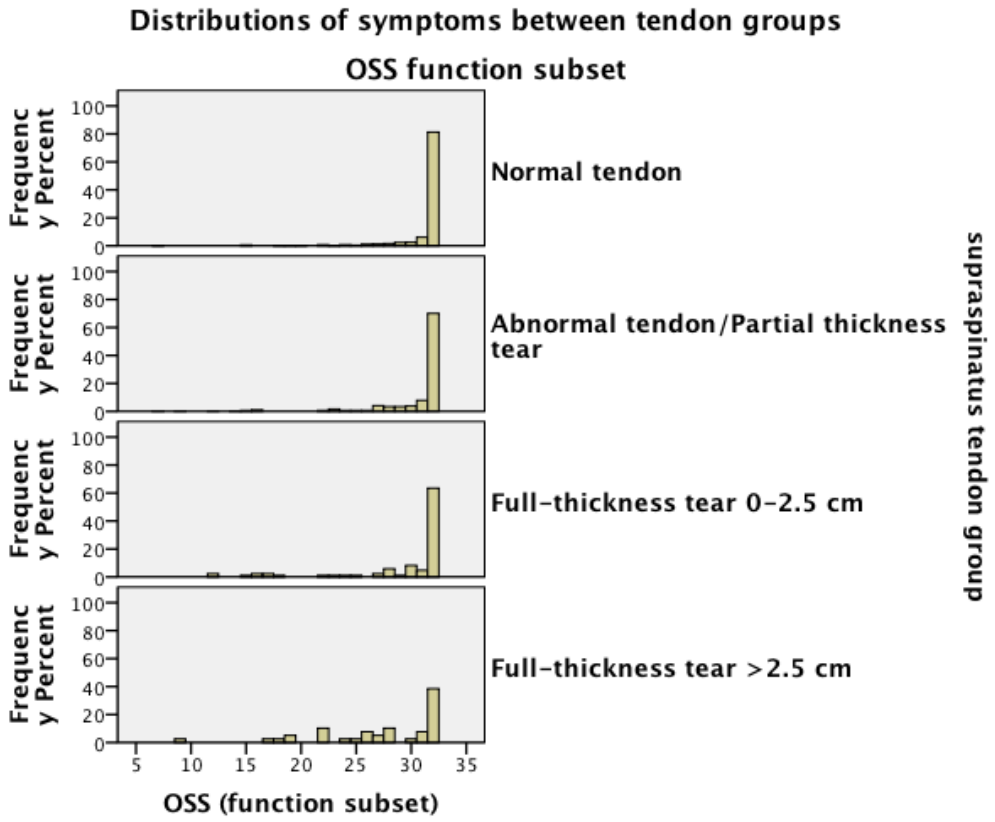
Additional problems are found when looking at the distributions of scores across different tear stage groups. There remains a negative skew in each of the groups. It is not simply that normal tendons consistently score well and larger tears consistently score worse, and thus the increased number of normal tendons is skewing the data. Also the variance of the distributions also differs between groups (see graphs 58-60). Thus, although non-parametric tests such as the Kruskal-Wallis test do not have any distributional assumptions and could be used to compare the groups, they do require that the distributions amongst the compared groups are the same, and therefore cannot be used with this data.



Graph 58: Distributions although all negatively skewed are not the same between tendon groups

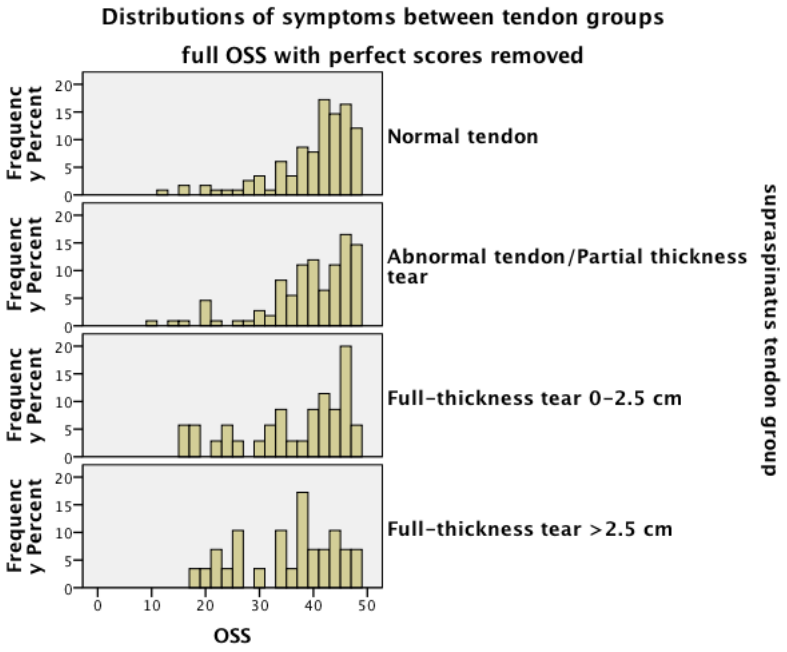


Graph 59: Distributions although all negatively skewed are not the same between tendon groups

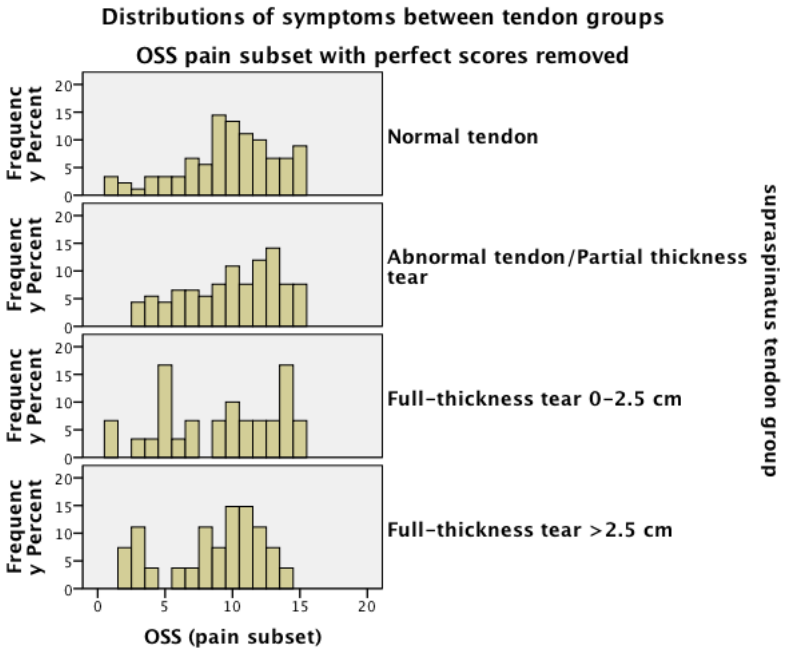


Graph 60: Distributions although all negatively skewed are not the same between tendon groups

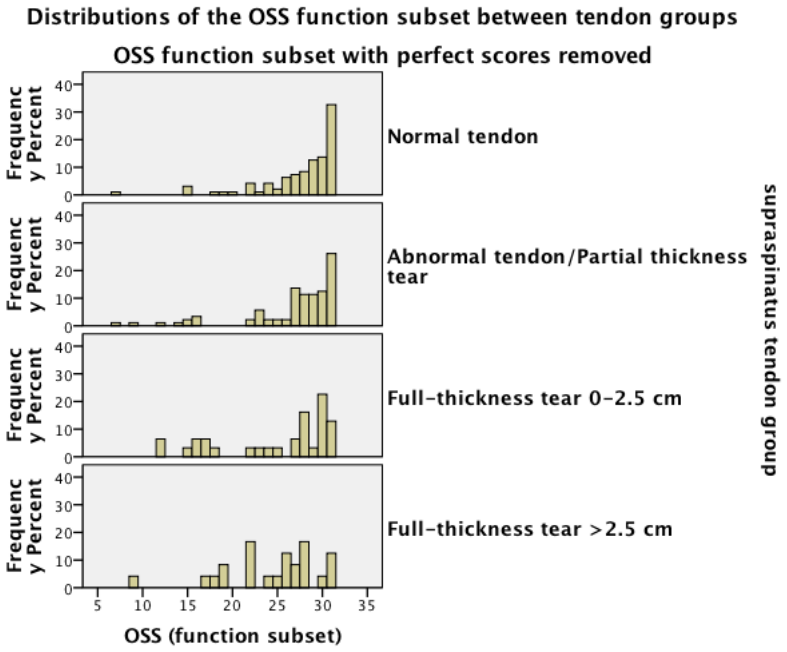
If however all perfect scores are removed from the data, leaving only symptomatic shoulders, non-parametric tests can be applied. Although the distributions are still not normal, the variance is now similar across each tear stage group and the assumptions of the Kruskal-Wallis test will hold (see graphs 61-63). It is important though that we are clear that the question being answered now only holds for symptomatic shoulders as different proportions of shoulders have been removed from different groups. Therefore symptom severity is only applicable to the cohort with symptoms and not the whole population.



Graph 61: With all perfect scores removed the variance between groups is now similar



Graph 62: With all perfect scores removed the variance between groups is now similar



Graph 63: With all perfect scores removed the variance between groups is now similar

REGRESSION ANALYSIS TO QUANTIFY DIFFERENCES IN SYMPTOM SEVERITY BETWEEN TENDON GROUPS

If there were a difference in symptom severity across the groups, a regression model would be required to quantify this. Creating a one-stage linear regression model using the complete scores is not valid due to the distribution of residuals caused by the heavily skewed data. Using the complete cohort would require a two-stage regression model. Here, the first stage would involve a regression model of perfect versus non-perfect scores, then a further regression mode would be fitted to the non-perfect scores and a combined co-efficient of determination produced. However, this model would be flawed for the following reasons:

1. Perfect tendon scores (normal tendons) do not predict perfect Oxford Shoulder Scores, thus it is inappropriate to split the data to fit the model accordingly.

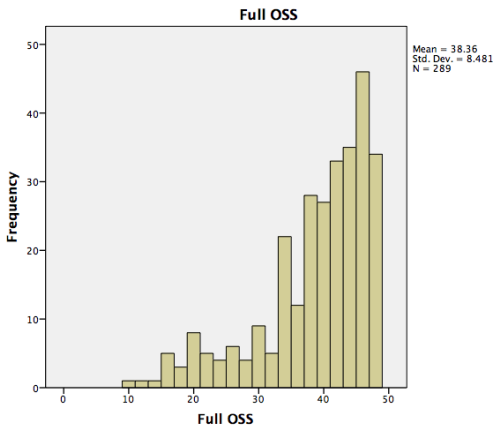
2. Tendon groups are a categorical variable and thus disproportionate numbers will be dropped from each group introducing bias into the remaining scores.

However, a one-stage regression model can be performed on all 'non-perfect score' data, but again giving information only about individuals with symptoms not the whole population in each group. In order to perform a linear regression model the OSS variables would need to be converted to a normal distribution.

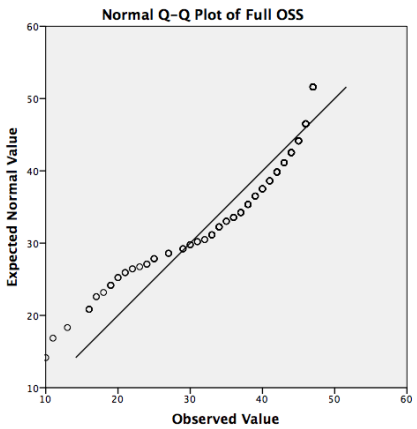
2 Attempts to normalise data

With all perfect scores removed the closest to normal distributions were achieved with a natural logarithmic transformation of the OSS, and OSS (function) once they had been inverted. The OSS (pain) was closest to a normal whilst untransformed (see graphs 77-88).

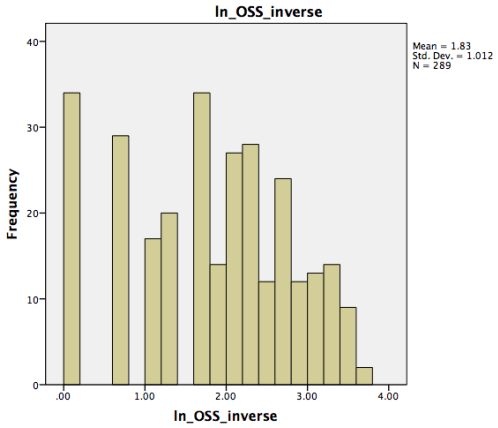
Full OSS



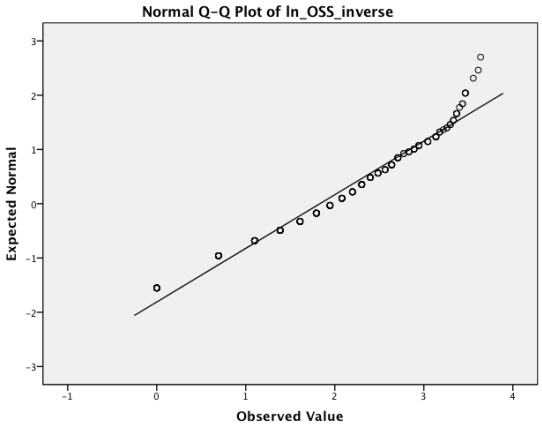
Graph 64: Negative skew of OSS



Graph 65: Abnormal distribution shown

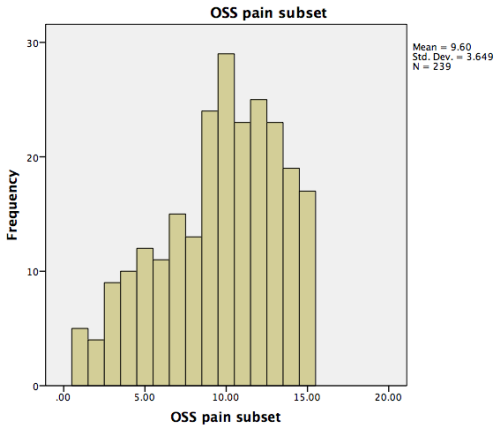


Graph 66: More normal distribution after transformation

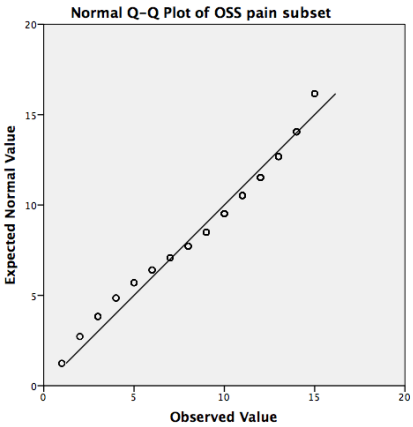


Graph 67: More normal Q-Q plot after transformation

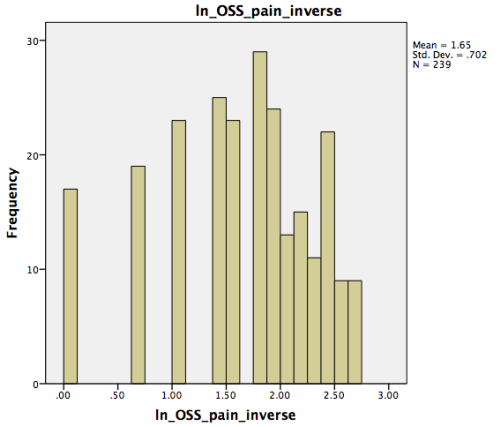
OSS Pain subset



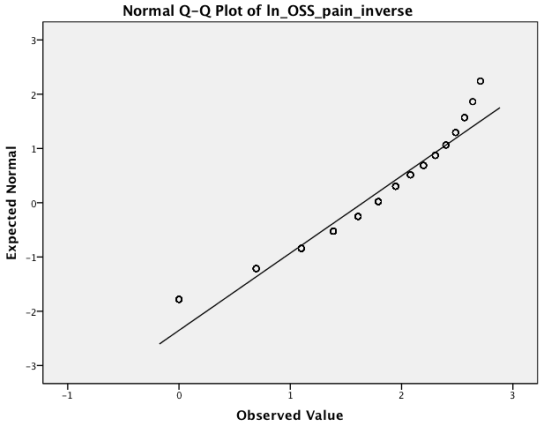
Graph 68: Near normal distribution of OSS pain



Graph 69: Normal distribution demonstrated

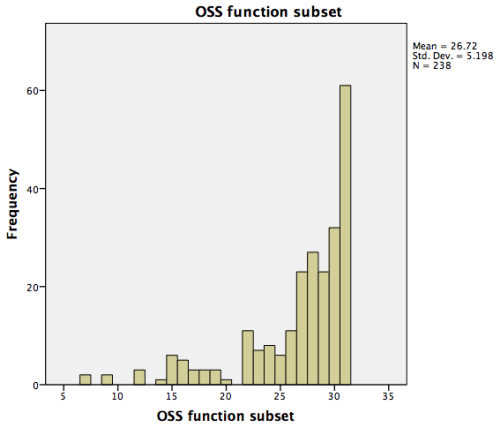


Graph 70: Distribution remains near normal

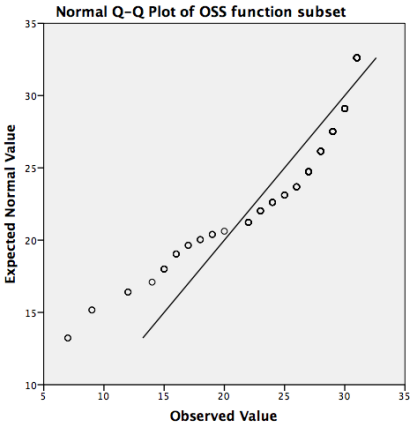


Graph 71: Similar Q-Q plot to pre-transformation

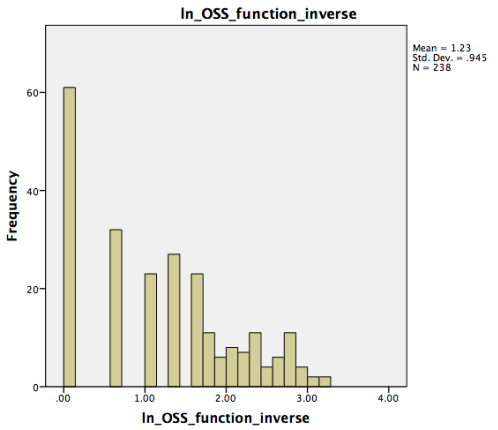
OSS Function subset



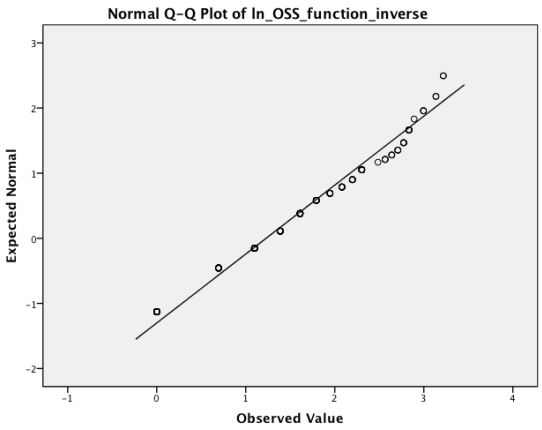
Graph 72: Negative skew of OSS function



Graph 73: Non normal distribution demonstrated



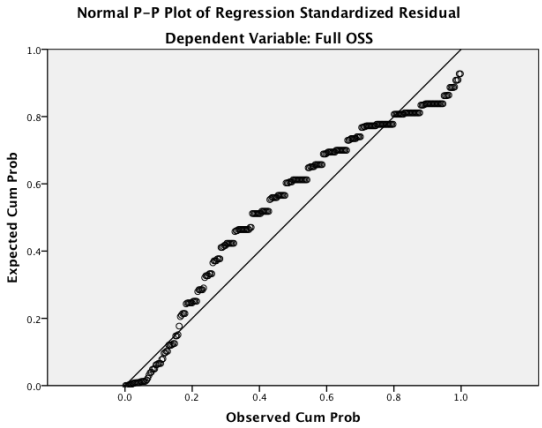
Graph 74: More normal distribution post transformation



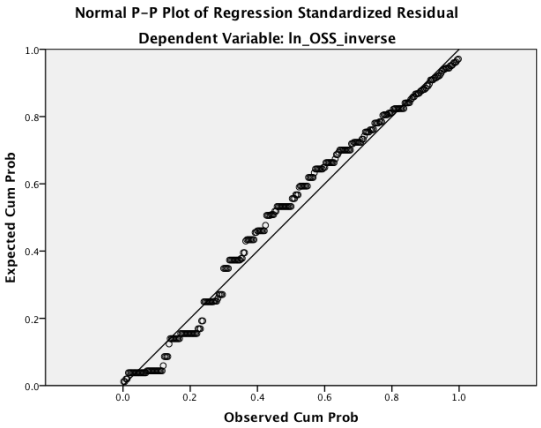
Graph 75: More normal Q-Q plot post transformation

To determine whether these transformations are adequate the planned regression models were performed (dependent variable: OSS; independent variable: Tear classification) using both the untransformed data and the transformed data, and the analysis of the residuals compared (see graphs 76-81). Here we can see that the residual analysis for the Ln (inverse OSS), Ln (inverse OSS function), performed better than the untransformed variables, and the OSS pain variable performed adequately without transformation.

OSS

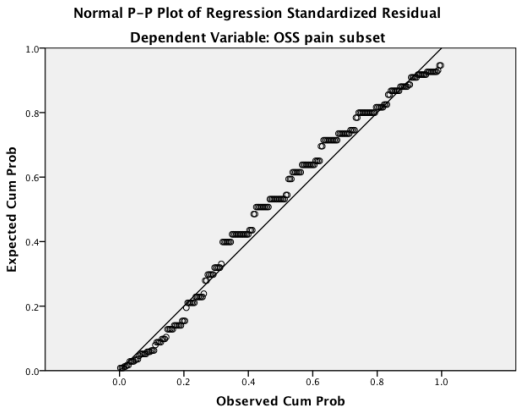


Graph 76: Abnormal residual plot with pre-transformed data

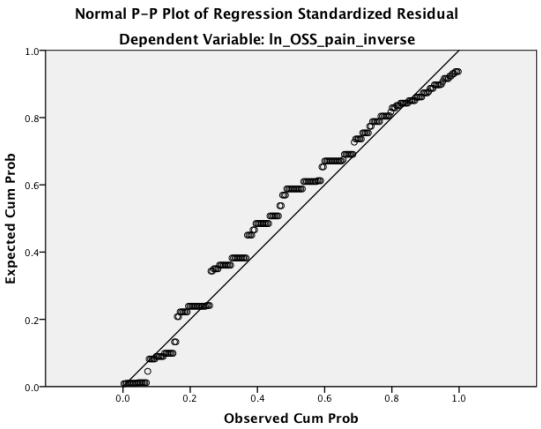


Graph 77: More normal residual plot post transformation

OSS Pain

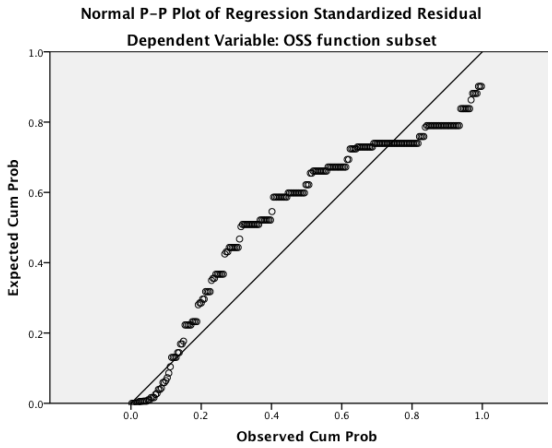


Graph 78: Adequate residual plot pre-transformation

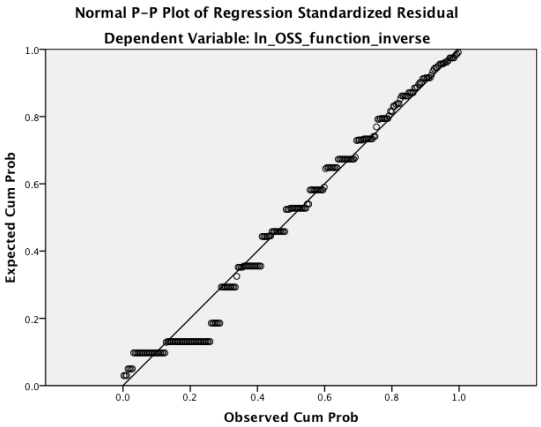


Graph 79: No improvement in residual plot post-transformation

OSS function



Graph 80: Abnormal residual plot pre-transformation



Graph 81: More normal residual plot post-transformation

Therefore, to answer whether the severity of symptoms is correlated to the severity of the pathology the following variables will be employed:

Ln (inverted OSS)

Ln (inverted OSS function)

OSS pain

3 Graphically representing data

For graphical representation of the data the OSS and OSS subscales have been divided into 4 groups. All perfect scores were labelled 'no symptoms', and then the remaining non–perfect scores were divided into 3 equal tertiles according to numbers of individuals. Statistical evaluation however will not be performed on these groups, but will take place on the continuous scale of each score. The score cut offs are laid out below in tables 50-52 and show the high proportion of normal scores:

OSS divided into 4 groups

	Frequency	Percent
No symptoms (48)	637	68.8
Mild symptoms (44-47)	100	10.8
Moderate symptoms (38-43)	94	10.2
Severe symptoms (0-37)	95	10.3
Total	926	100.0

Table 50

OSS pain subset divided into 4 groups

	Frequency	Percent
No symptoms (16)	687	74.2
Mild symptoms (12-15)	84	9.1
Moderate symptoms (9-11)	76	8.2
Severe symptoms (0-8)	79	8.5
Total	926	100.0

Table 51

OSS function subset divided into 4 groups

	Frequency	Percent
No symptoms (32)	688	74.3
Mild symptoms (30-31)	93	10.0
Moderate symptoms (27-29)	73	7.9
Severe symptoms (0-26)	72	7.8
Total	926	100.0

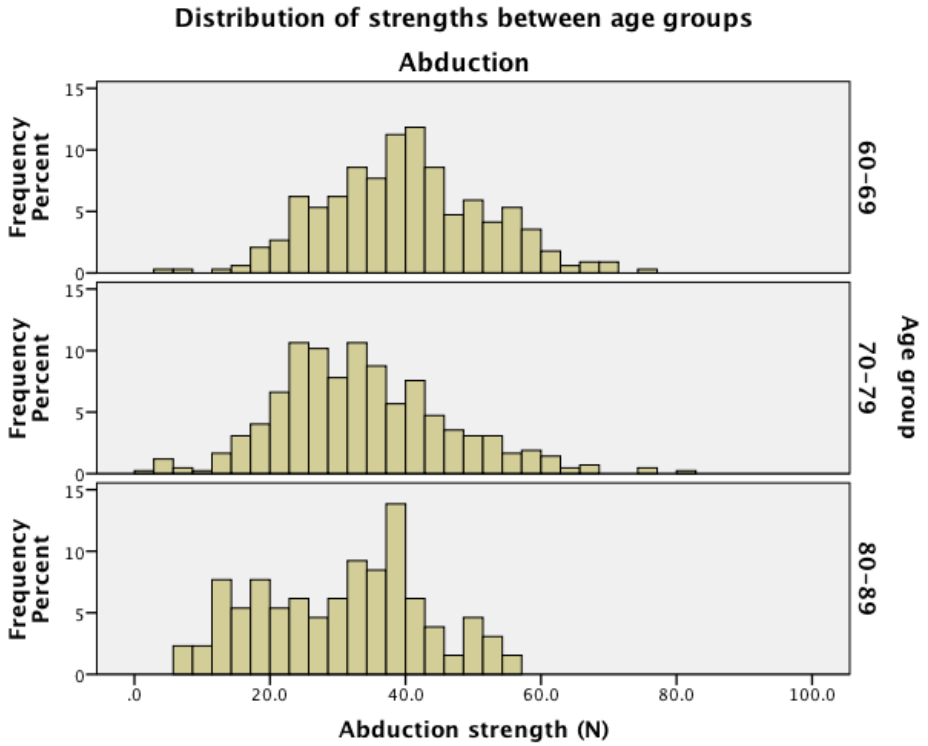
Table 52

Appendix 2: Motor strength testing; establishment of potential confounders

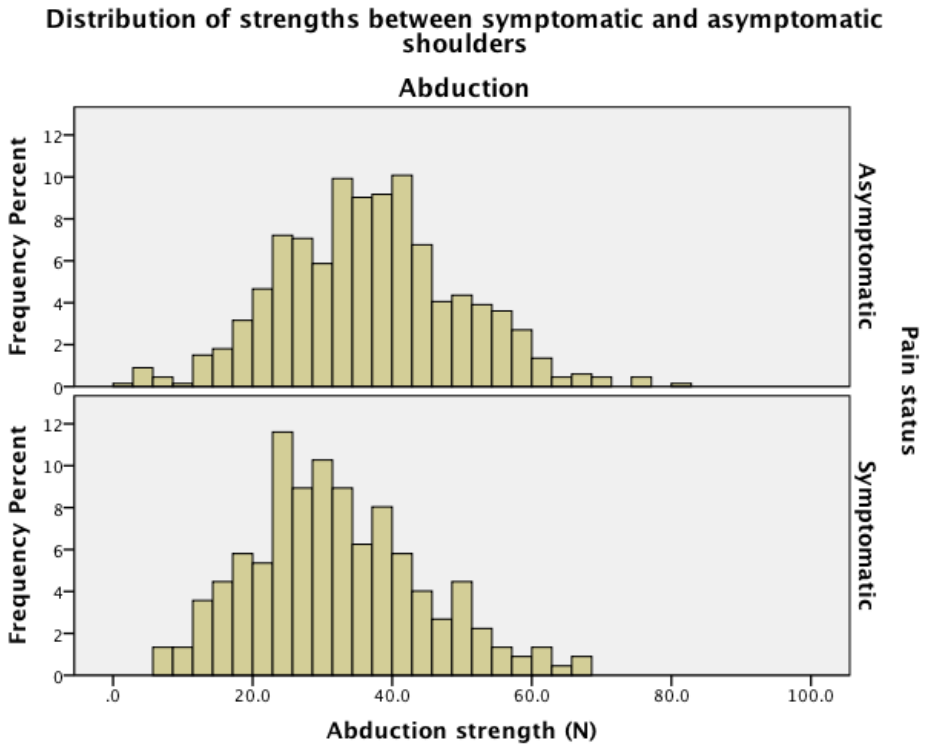
An aim of this thesis is to determine whether rotator cuff tears are associated with functional deficits, and whether these are quantifiable, rather than being solely subjective. Strength of the shoulder girdle however, will be affected by a number of external factors, not only the rotator cuff function. Therefore any statistical model that uses rotator cuff pathology to predict strength will need to take into account additional factors. This sub study will investigate the role of age, BMI, arm dominance, and pain as potential confounders to the primary question.

1 Descriptive statistics of potential confounders

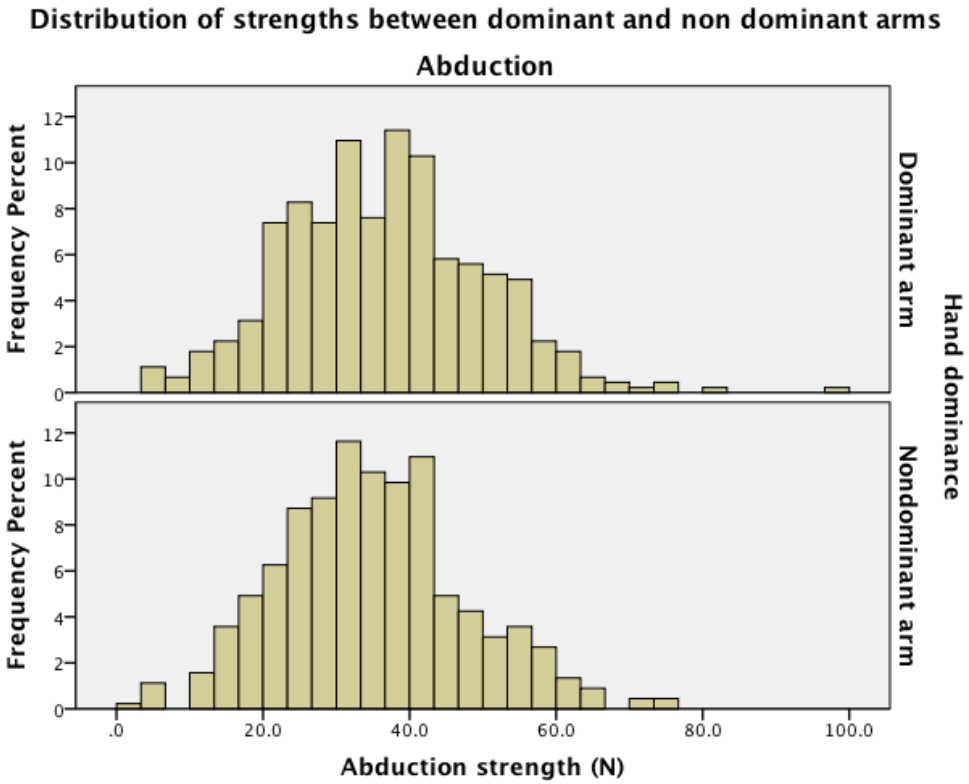
Mean shoulder abduction strengths were compared across different age and BMI groups. They were also compared between dominant and non-dominant arms and painful and pain free arms (see graphs 82-85 and tables 53-56).



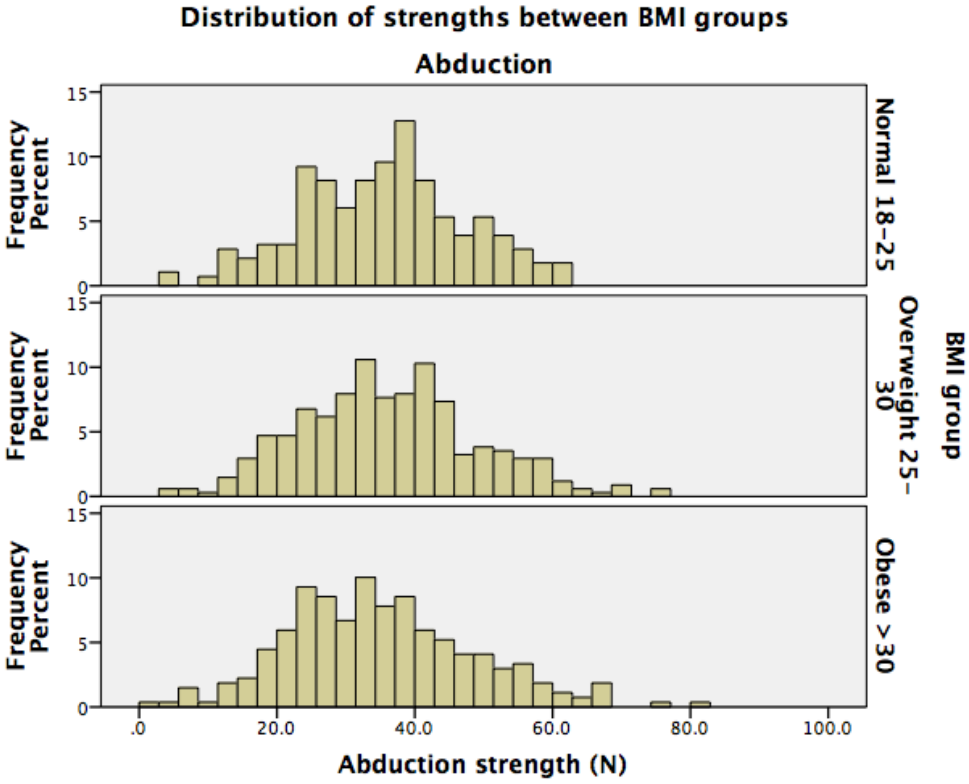
Graph 82: Mean shoulder abduction strength decreased which each decile of age



Graph 83: Mean shoulder abduction strength decreased in the presence of shoulder symptoms



Graph 84: Mean shoulder abduction strength not significantly different between the dominant and non-dominant arms



Graph 85: Mean shoulder abduction strength was not affected by BMI

Mean abduction strength according to age group

Age group (n=891)	n	Mean	95% confidence intervals	
			Lower	Upper
60-69	338	39.872	38.609	41.135
70-79	423	33.643	32.426	34.861
80-89	130	30.457	28.309	32.606

Table 53: Mean shoulder abduction strength decreased which each decile of age

Mean abduction strength according to symptoms

Pain status (n=891)	n	Mean	95% confidence intervals	
			Lower	Upper
Asymptomatic	665	36.798	35.827	37.770
Symptomatic	224	31.857	30.230	33.484

Table 54: Mean shoulder abduction strength decreased in the presence of shoulder symptoms

Mean abduction strength according to arm dominance

Hand dominance	n	Mean	95% confidence intervals	
			Lower	Upper
Dominant	447	36.470	35.222	37.719
Non dominant	447	34.906	33.727	36.086

Table 55: Mean shoulder abduction strength not significantly different between the dominant and non-dominant arms

Mean abduction strength according to BMI

BMI group (n=891)	n	Mean	95% confidence intervals	
			Lower	Upper
BMI 18-25	282	35.328	33.928	36.728
BMI 25-30	340	36.128	34.747	37.509
BMI >30	269	35.023	33.394	36.651

Table 56: Mean shoulder abduction strength was not affected by BMI

2 Analysis of potential confounders

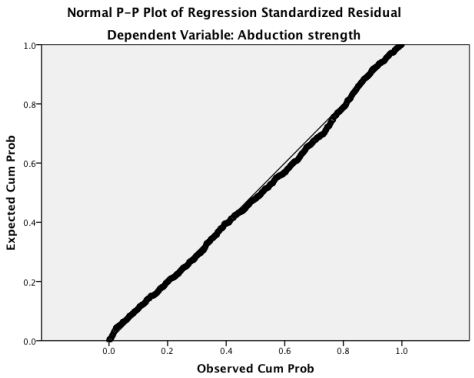
One way ANOVA tests showed there to be a significant difference in shoulder abduction strength according to age group ($p<0.001$), but no difference according to BMI group ($p=0.541$). Independent T-tests showed that pain free shoulders were significantly stronger than painful shoulders (4.94N, $p<0.001$) and paired T-tests showed dominant arms were significantly stronger than non-dominant arms (difference 1.56N, $p<0.001$). (Independent T-test showed no difference (difference=1.56, $p=0.074$).

Linear regression models showed that age, and pain status had an overall effect on strength, but BMI or whether the shoulder was the dominant or non-dominant arm had no statistically significant effect (For regression analysis see tables 57-61, with residual analysis, graphs 86-90).

Regression co-efficients - Age

	Unstandardized Coefficients		95.0% Confidence Interval for B	
	B	Sig.	Lower Bound	Upper Bound
Constant	84.464	.000	73.840	95.088
Age	-0.678	.000	-0.825	-0.531

Table 57: Significant effect of age on shoulder abduction strength

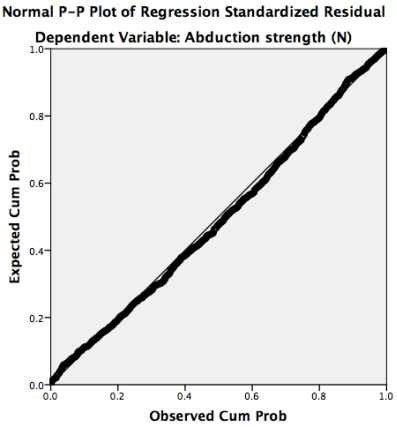


Graph 86: Residual analysis – Normal Q-Q plot demonstrating a well performing regression model

Regression co-efficients – Age groups

	Unstandardized Coefficients B	Sig.	95.0% Confidence Interval for B	
			Lower Bound	Upper Bound
60-69 (Constant)	39.872	.000	11.765	22.923
70-79	-6.229	.000	-7.996	-4.462
80-89	-9.415	.000	-11.914	-6.915

Table 58: Significant effect of age group on shoulder abduction strength

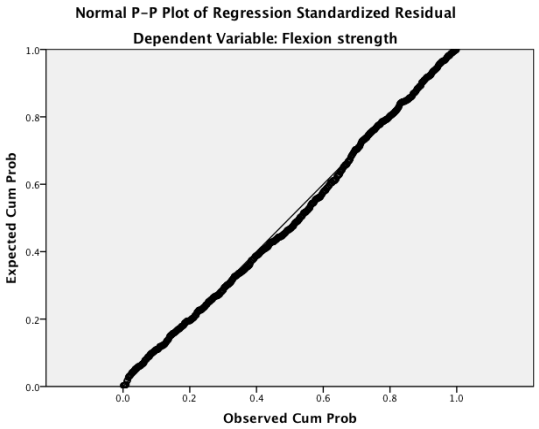


Graph 87: Residual analysis – Normal Q-Q plot demonstrating a well performing regression model

Regression co-efficients – Symptoms

	Unstandardized Coefficients B	Sig.	95.0% Confidence Interval for B	
			Lower Bound	Upper Bound
Asymptomatic (Constant)	41.510	.000	40.484	42.536
Symptomatic	-6.287	.000	-8.330	-4.243

Table 59: Significant effect of symptoms on shoulder abduction strength

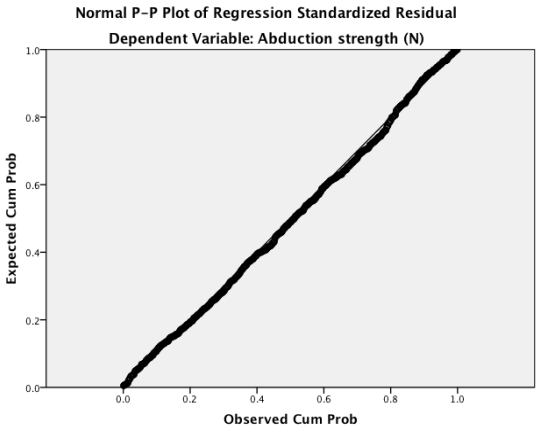


Graph 88: Residual analysis – Normal Q-Q plot demonstrating a well performing regression model

Regression co-efficients – Arm dominance

	Unstandardized Coefficients B	Sig.	95.0% Confidence Interval for B	
			Lower Bound	Upper Bound
Dominant (Constant)	37.907	.000	35.245	40.570
Non dominant	-1.578	.066	-3.262	.107

Table 60: No significant effect of arm dominance on shoulder abduction strength

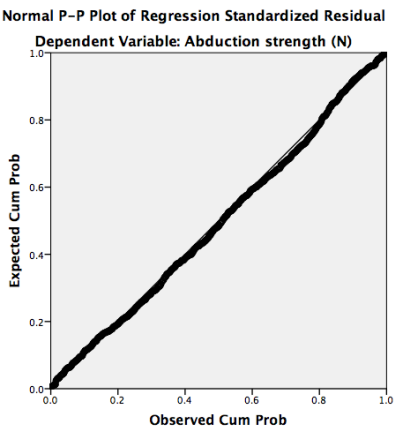


Graph 89: Residual analysis – Normal Q-Q plot demonstrating a well performing regression model

Regression co-efficients – BMI

	Unstandardized Coefficients B	Sig.	95.0% Confidence Interval for B	
			Lower Bound	Upper Bound
Constant	34.964	.000	30.233	39.694
BMI	0.021	.808	-0.146	0.188

Table 61: No significant effect of BMI on shoulder abduction strength



Graph 90: Residual analysis – Normal Q-Q plot demonstrating a well performing regression model

ANALYSIS OF INTERACTIONS BETWEEN TENDONOPATHY GROUPS AND CONFOUNDERS

The above analysis has suggested that along with rotator cuff pathology other factors may also be predictors of shoulder strength. However, chapters one and two have shown that rotator cuff tears are associated with an increased prevalence of shoulder pain. Tears are also more prevalent in the older age groups and the dominant arm. Therefore it is not possible using uni-variable models to determine which variables are the main contributors to loss of strength. For example, if there were a decrease in strength with rotator cuff tear, one would expect to see a decrease in strength with age and painful shoulders as tears are more prevalent in the older age groups, and are associated with pain. Inherently one would expect the dominant arm to be stronger than the non-dominant arm, but above has shown there to be no difference. This could be also due to the increased prevalence of rotator cuff tear in the dominant arm. However, it could also be that age and pain are in fact more accountable for loss of strength than rotator cuff tears, and these are confounding our initial model introducing bias of association.

In order to determine the true effects of each variable a multivariable model must be used which will determine, individual effect and any interacting effects. The graphs and tables below show the interactions and confounding effects of the variables, and the build of the optimum model for analysis.

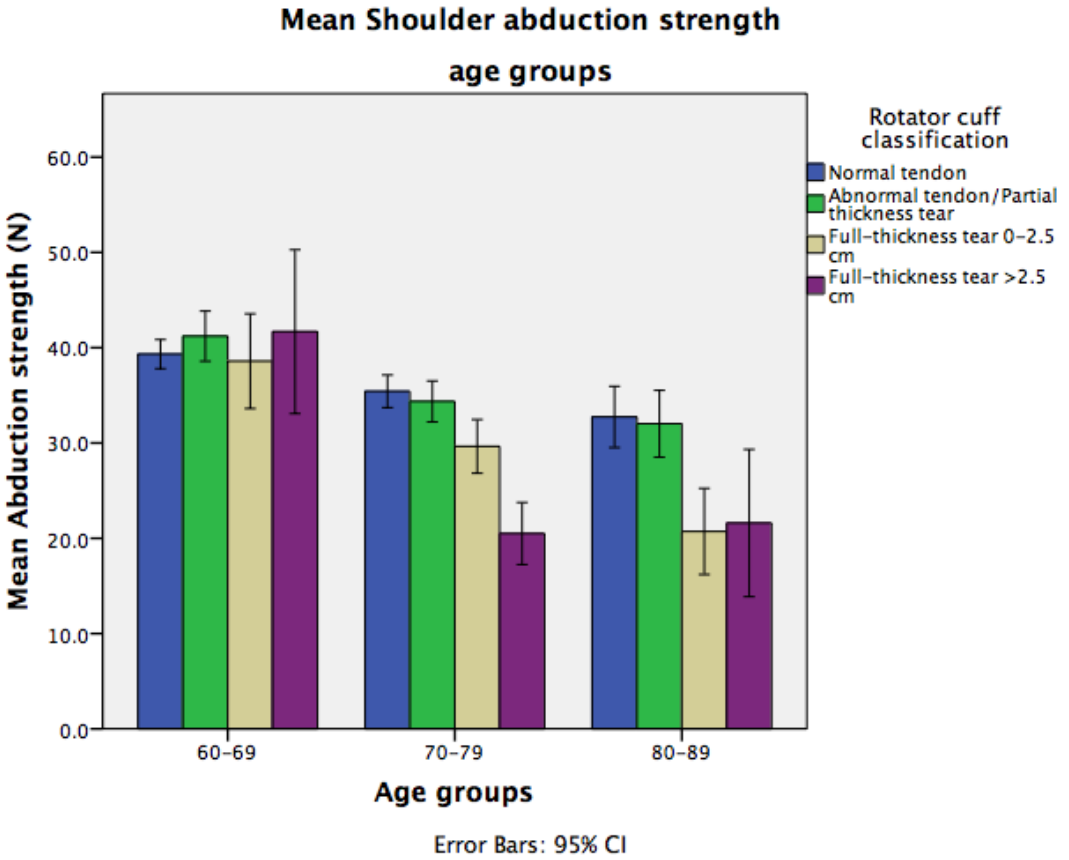
Rotator cuff tendinopathy and age

Age appears to have both a confounding and an interacting effect (tables 62 & 63, graph 91).

Mean abduction strength according to age and tendinopathy

	60-69	70-79	80-89	Total
Normal tendon	39.3	35.4	32.7	36.8
Abnormal /Partial tear	41.2	34.4	32.0	36.3
Full-thickness tear 0-2.5 cm	38.6	29.6	20.7	30.4
Full-thickness tear >2.5 cm	41.7	20.5	21.6	24.1
Total	39.9	33.6	30.5	35.5

Table 62: Mean abduction strength was maintained irrespective of rotator cuff tear in the 60-69 age group but in the >70's decreased in the presence of rotator cuff tear demonstrating an interaction between tendinopathy and age



Graph 91: Mean abduction strength was maintained irrespective of rotator cuff tear in the 60-69 age group but in the >70's decreased in the presence of rotator cuff tear demonstrating an interaction between tendinopathy and age

Regression model: effect of co-variates: Tendinopathy and age

	df	Sig.	Observed Power
Corrected Model	11	.000	1.000
Intercept	1	.000	1.000
Tendon classification	3	.000	.993
Age group	2	.000	1.000
Tendon classification/ age group interaction	6	.011	.881

Table 63: Both tendinopathy group and age had an independent effect on strength but also did the interaction between the two

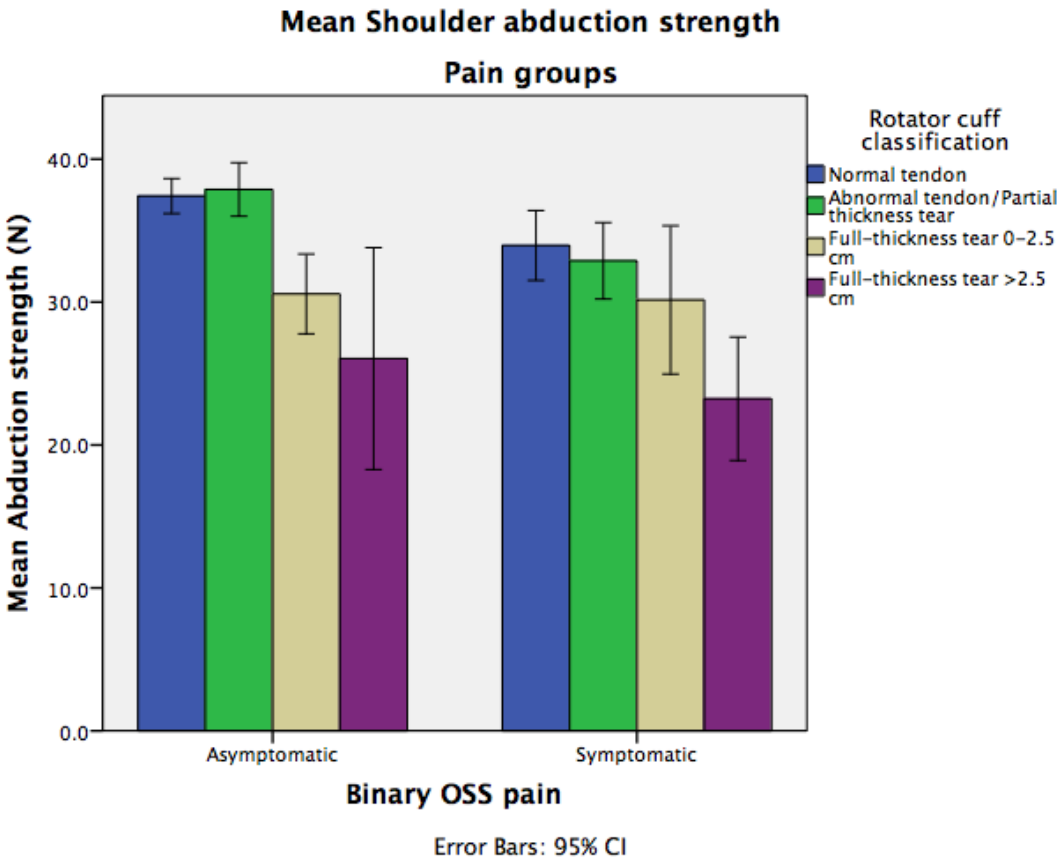
Rotator cuff tendinopathy and pain

Pain has a confounding effect but no obvious interactions. (Tables 64-65, Graph 92)

Mean abduction strength according to symptoms and tendinopathy

	Asymptomatic	Symptomatic	Total
Normal tendon	37.4	34.0	36.8
Abnormal /Partial tear	37.9	32.9	36.3
Full-thickness tear 0-2.5 cm	30.6	30.2	30.4
Full-thickness tear >2.5 cm	26.0	23.2	24.1
Total	36.8	31.9	35.6

Table 64: Strength was uniformly decreased in the symptomatic shoulders to the same extent across each tendinopathy



Graph 92: Strength was uniformly decreased in the symptomatic shoulders to the same extent across each tendinopathy

Regression model: Effect of co-variates: tendinopathy and pain

Model 1	df	Sig.	Observed Power
Corrected Model	7	.000	1.000
Intercept	1	.000	1.000
Tendon classification	3	.000	.999
Pain status	1	.040	.538
Tendon classification/ pain status interaction	3	.571	.192

Model 2 (multi-variable)	df	Sig.	Observed Power
Corrected Model	12	.000	1.000
Intercept	1	.000	1.000
Tendon classification	3	.000	.966
Age group	2	.000	1.000
Pain status	1	.000	.997
Tendon classification/ age group interaction	6	.008	.898

Table 65 Pain and tendinopathy group had an effect on strength with no interaction between the two. Age and its interaction remained significant when added to the model

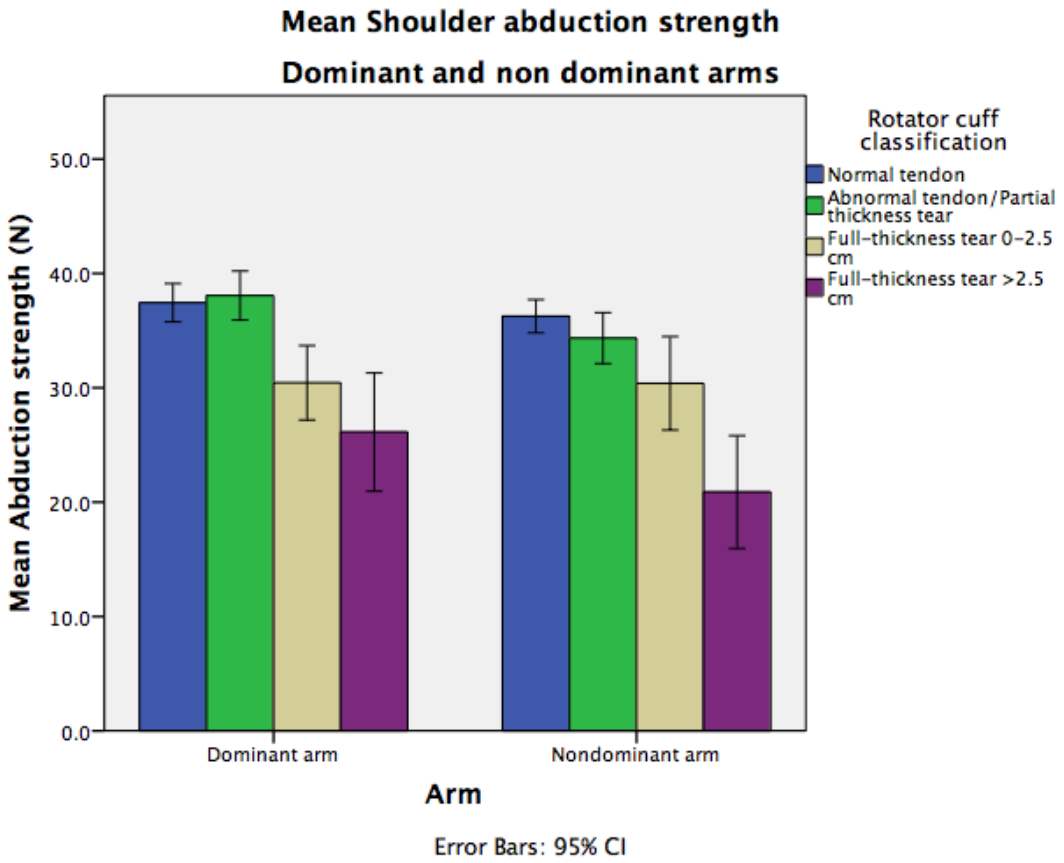
Rotator cuff tendinopathy and arm dominance

In a model alone whether it was a dominant or non-dominant arm had no confounding effect on shoulder abduction strength. However, when broken down into tendinopathy groups, it is apparent that the dominant arm was stronger universaly across all groups, and actually does have a confounding effect on strength. This is shown in tables 66-67 and graph 93. The confounding effect only becomes apparent once in a combined model, as on its own it is being masked by the greater proportion of dominant arms having a FTT which have associated less strength.

Mean abduction strength according to arm dominance and tendinopathy

	Dominant	Non-Dominant	Total
Normal tendon	37.4	36.3	36.8
Abnormal /Partial tear	38.1	34.3	36.3
Full-thickness tear 0-2.5 cm	30.4	30.4	30.4
Full-thickness tear >2.5 cm	26.1	20.9	24.1
Total	36.3	34.8	35.5

Table 66: Mean abduction strengths were universally lower in the non-dominant shoulder across all tendon groups



Graph 93: Mean abduction strengths were universally lower in the non-dominant shoulder across all tendon groups

Regression model: Effect of covariates: tendinopathy & arm dominance

Model 1	df	Sig.	Observed Power
Corrected Model	7	.000	1.000
Intercept	1	.000	1.000
Tendon classification	3	.000	1.000
Dominant/Non-dominant	1	.060	.470
Tendon classification/ dominance interaction	3	.400	.269

Model 2	df	Sig.	Observed Power
Corrected Model	13	.000	1.000
Intercept	1	.000	1.000
Tendon classification	3	.000	.974
Age group	2	.000	1.000
Pain status	1	.000	.997
Hand dominance	1	.011	.716
Tendon classification/ age group interaction	6	.009	.891

Table 67 Regression model showed that tendinopathy group, age, pain, arm dominance, and an interaction between age and tendinopathy group, all were associated with changes in shoulder abduction strength. This though was only apparent in a combined multi-variate model

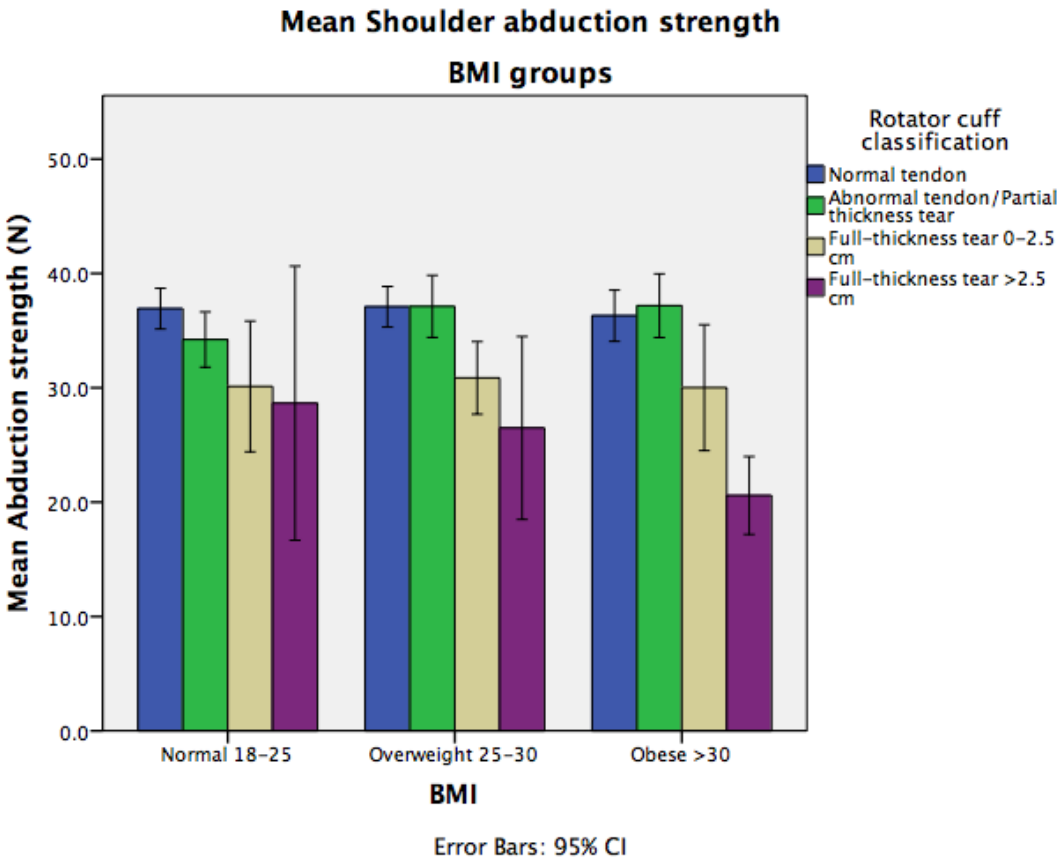
Rotator cuff tendinopathy and BMI

BMI had no confounding or interacting effect in any model (table 68 and graph 94).

Mean abduction strength according to BMI and tendinopathy

	Normal 18-25	Overweight 25-30	Obese >30	Total
Normal tendon	36.9	37.1	36.3	36.8
Abnormal /Partial tear	34.2	37.1	37.2	36.3
Full-thickness tear 0-2.5 cm	30.1	30.9	30.0	30.4
Full-thickness tear >2.5 cm	28.7	26.5	20.6	24.1
Total	35.3	36.1	35.0	35.5

Table 68: Mean abduction strength was no different between BMI groups



Graph 94: Mean shoulder abduction strength was no different between BMI groups

OVERALL BEST MODEL

The best model at predicting strength included rotator cuff tendinopathy classification, age, pain status, arm dominance, and the interaction between age and rotator cuff tendinopathy.

Bibliography

1. Urwin M, Symmons D, Allison T, Brammah T, Busby H, Roxby M, et al. Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. Ann Rheum Dis. 1999 ed. 1998 Nov;57(11):649–55.

2. Pope DP, Croft PR, Pritchard CM, Silman AJ. Prevalence of shoulder pain in the community: the influence of case definition. Ann Rheum Dis. 1997 May;56(5):308–12.

3. RJ M, JW S, TP S. Persistent shoulder pain: epidemiology, pathophysiology, and diagnosis. Am J Orthop (Belle Mead NJ). 2005 Dec 1;34(12 Suppl):5–9.

4. Roquelaure Y, Ha C, Leclerc A, Touranchet A, Sauteron M, Melchior M, et al. Epidemiologic surveillance of upper-extremity musculoskeletal disorders in the working population. Arthritis Rheum. 2006;55(5):765–78.

5. Arden N. The association between osteoarthritis and osteoporotic fracture: the Chingford Study. Rheumatology. Arthritis and Rheumatology; 1996 Dec 1;35(12):1299–304.

6. Hart DJ, Mootoosamy I, Doyle DV, Spector TD. The relationship between osteoarthritis and osteoporosis in the general population: the Chingford Study. Ann Rheum Dis. 1994 Mar 1;53(3):158–62.

7. DJ H, TD S. The relationship of obesity, fat distribution and osteoarthritis in women in the general population: the Chingford Study. J Rheumatol. 1993 Feb 1;20(2):331–5.

8. De Jesus JO, Parker L, Frangos AJ, Nazarian LN. Accuracy of MRI, MR Arthrography, and Ultrasound in the Diagnosis of Rotator Cuff Tears: A Meta-Analysis. American Journal of Roentgenology. 2009 May 19;192(6):1701–7.

9. Naqvi GA, Jadaan M, Harrington P. Accuracy of ultrasonography and magnetic resonance imaging for detection of full thickness rotator cuff tears. Int J Shoulder Surg. 2009 Oct;3(4):94–7.

10. Dinnes J, Loveman E, McIntyre L, Waugh N. The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review. Health Technol Assess. 2003rd ed. 2003;7(29):iii–1–166.

11. Teefey SA, Rubin DA, Middleton WD, Hildebolt CF, Leibold RA, Yamaguchi K. Detection and Quantification of Rotator Cuff TearsComparison of Ultrasonographic, Magnetic Resonance Imaging, and Arthroscopic

Findings in Seventy-one Consecutive Cases. J Bone Joint Surg Am. The Journal of Bone and Joint Surgery; 2004 Apr 1;86(4):708–16.

12. Ottenheijm RP, Jansen MJ, Staal JB, van den Bruel A, Weijers RE, de Bie RA, et al. Accuracy of diagnostic ultrasound in patients with suspected subacromial disorders: a systematic review and meta-analysis. Arch Phys Med Rehabil. 2010 ed. 91(10):1616–25.

13. Smith TO, Back T, Toms AP, Hing CB. Diagnostic accuracy of ultrasound for rotator cuff tears in adults: a systematic review and meta-analysis. Clin Radiol. 2011 Nov;66(11):1036–48.

14. Teefey SA, Hasan SA, Middleton WD, Patel M, Wright RW, Yamaguchi K. Ultrasonography of the rotator cuff. A comparison of ultrasonographic and arthroscopic findings in one hundred consecutive cases. J Bone Joint Surg Am. 2000 ed. 2000 Apr;82(4):498–504.

15. Teefey SA, Middleton WD, Payne WT, Yamaguchi K. Detection and measurement of rotator cuff tears with sonography: analysis of diagnostic errors. American Journal of Roentgenology. 2005 Jun;184(6):1768–73.

16. Middleton WD, Teefey SA, Yamaguchi K. Sonography of the rotator cuff: analysis of interobserver variability. American Journal of Roentgenology. 2004 Nov;183(5):1465–8.

17. Iannotti JP, Ciccone J, Buss DD, Visotsky JL, Mascha E, Cotman K, et al. Accuracy of office-based ultrasonography of the shoulder for the diagnosis of rotator cuff tears. J Bone Joint Surg Am. 2005 Jun;87(6):1305–11.

18. Al-Shawi AK, Badge R, Bunker TD. Reliability of shoulder ultrasound to detect rotator cuff tears when performed by the surgeon in the place of first contact. J Bone Joint Surg: Orthopaedic Proceedings 2012 Feb.

19. Moosmayer S, Smith H-J. Diagnostic ultrasound of the shoulder--a method for experts only? Results from an orthopedic surgeon with relative inexpensive compared to operative findings. Acta Orthop. 2005 Aug;76(4):503–8.

20. Jeyam M, Funk L, Harris J. Are shoulder surgeons any good at diagnosing rotator cuff tears using ultrasound?: A comparative analysis of surgeon vs radiologist. Int J Shoulder Surg. 2008 ed. 2008 Jan;2(1):4–6.

21. Murphy RJ, Daines MT, Carr AJ, Rees JL. An independent learning method for orthopaedic surgeons performing shoulder ultrasound to identify full-thickness tears of the rotator cuff. J Bone Joint Surg Am. 2013 Feb 6;95(3):266–72.

22. O'Neill JMD. Musculoskeletal Ultrasound. Springer; 2008.

23. Ruotolo C, Fow JE, Nottage WM. The supraspinatus footprint: an anatomic study of the supraspinatus insertion. Arthroscopy. 2004 Mar;20(3):246–9.

24. Roh MS, Wang VM, April EW, Pollock RG, Bigliani LU, Flatow EL. Anterior and posterior musculotendinous anatomy of the supraspinatus. J Shoulder Elbow Surg. 2000 Sep;9(5):436–40.

25. Minagawa H, Itoi E, Konno N, Kido T, Sano A, Urayama M, et al. Humeral attachment of the supraspinatus and infraspinatus tendons: an anatomic study. Arthroscopy. 1998 Apr;14(3):302–6.

26. Mochizuki T, Sugaya H, Uomizu M, Maeda K, Matsuki K, Sekiya I, et al. Humeral insertion of the supraspinatus and infraspinatus. New anatomical findings regarding the footprint of the rotator cuff. Surgical technique. J Bone Joint Surg Am. 2009 Mar 1;91 Suppl 2 Pt 1:1–7.

27. Codman EA, Chipman WW, Clark JG, Kanavel AB. Standardisation of hospitals: report of the committee appointed by the Clinical Congress of Surgeons of North America. Trans Clin Cong Surg North Am 1913;4:2-8.

28. Harvie P, Pollard TCB, Chennagiri RJ, Carr AJ. The use of outcome scores in surgery of the shoulder. J Bone Joint Surg Br. 2005 Feb;87(2):151–4.

29. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. Clinical Orthopaedics and Related Research. 1987 ed. 1987 Jan;(214):160–4.

30. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. J Bone Joint Surg Br. 1996 ed. 1996 Jul;78(4):593–600.

31. Dawson J, Rogers K, Fitzpatrick R, Carr A. The Oxford shoulder score revisited. Arch Orthop Trauma Surg. 2008 ed. 2009 Jan;129(1):119–23.

32. Richards RR, An K-N, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG, et al. A standardized method for the assessment of shoulder function. Journal of Shoulder and Elbow Surgery. 1994 Nov;3(6):347–52.

33. Dawson J, Hill G, Fitzpatrick R, Carr A. Comparison of clinical and patient-based measures to assess medium-term outcomes following shoulder surgery for disorders of the rotator cuff. Arthritis Rheum. 2002 Oct 15;47(5):513–9.

34. Conboy VB, Morris RW, Kiss J, Carr AJ. An evaluation of the Constant-Murley shoulder assessment. J Bone Joint Surg Br. 1996 Mar;78(2):229–32.

35. Wright JG, Rudicel S, Feinstein AR. Ask patients what they want. Evaluation of individual complaints before total hip replacement. J Bone Joint Surg Br. 1994 Mar;76(2):229–34.

36. Garratt A. Quality of life measurement: bibliographic study of patient assessed health outcome measures. BMJ. British Medical Journal Publishing Group; 2002 Jun 15;324(7351):1417–7.

37. Paul A, Lewis M, Shadforth MF, Croft PR, Van Der Windt DAWM, Hay EM. A comparison of four shoulder-specific questionnaires in primary care. Ann Rheum Dis. 2004 Oct;63(10):1293–9.

38. Olley LM, Carr AJ. The use of a patient-based questionnaire (the Oxford Shoulder Score) to assess outcome after rotator cuff repair. Ann R Coll Surg Engl. 2008 ed. 2008 May;90(4):326–31.

39. Cloke DJ, Lynn SE, Watson H, Steen IN, Purdy S, Williams JR. A comparison of functional, patient-based scores in subacromial impingement. J Shoulder Elbow Surg. 2005 Jul;14(4):380–4.

40. Ekeberg OM, Bautz-Holter E, Tveitå EK, Keller A, Juel NG, Brox JI. Agreement, reliability and validity in 3 shoulder questionnaires in patients with rotator cuff disease. BMC Musculoskelet Disord. 2008;9:68.

41. Younis F, Sultan J, Dix S, Hughes PJ. The range of the Oxford Shoulder Score in the asymptomatic population: a marker for post-operative improvement. Ann R Coll Surg Engl. 2011 Nov;93(8):629–33.

42. Croft P, Pope D, Zonca M, O'Neill T, Silman A. Measurement of shoulder related disability: results of a validation study. Ann Rheum Dis. 1994 ed. 1994 Aug;53(8):525–8.

43. Roy J-S, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic review of four questionnaires. Arthritis Rheum. 2009 May 15;61(5):623–32.

44. Briggs M, Closs JS. A descriptive study of the use of visual analogue scales and verbal rating scales for the assessment of postoperative pain in orthopedic patients. J Pain Symptom Manage. 1999 Dec;18(6):438–46.

45. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. J Clin Nurs. 2005 Aug;14(7):798–804.

46. Lara-Muñoz C, De Leon SP, Feinstein AR, Puente A, Wells CK. Comparison of three rating scales for measuring subjective phenomena in clinical research. I. Use of experimentally controlled auditory stimuli. Arch Med Res. 2004 Jan;35(1):43–8.

47. Chibnall JT, Tait RC. Pain assessment in cognitively impaired and unimpaired older adults: a comparison of four scales. Pain. 2001 May;92(1-2):173–86.

48. Miranda H, Gold JE, Gore R, Punnett L. Recall of prior musculoskeletal pain. Scand J Work Environ Health. 2006 Aug;32(4):294–9.

49. Broderick JE, Stone AA, Calvanese P, Schwartz JE, Turk DC. Recalled pain ratings: a complex and poorly defined task. J Pain. 2006 Feb;7(2):142–9.

50. Beard DJ, Carr AJ, Dawson J. The use of the Oxford hip and knee scores. J

Bone Joint Surg Br. 2007;89-B:1010-14.

51. Judge A, Arden NK, Price A, Glyn-Jones S, Beard D, Carr AJ, et al. Assessing patients for joint replacement: can pre-operative Oxford hip and knee scores be used to predict patient satisfaction following joint replacement surgery and to guide patient selection? J Bone Joint Surg Br. 2011 Dec;93(12):1660–4.

52. Kalairajah Y, Azurza K, Hulme C, Molloy S, Drabu KJ. Health outcome measures in the evaluation of total hip arthroplasties--a comparison between the Harris hip score and the Oxford hip score. J Arthroplasty. 2005 Dec;20(8):1037–41.

53. Van Kampen DA, Willems WJ. Determination and comparison of the smallest detectable change (SDC) and the minimal important change (MIC) of four-shoulder patient-reported outcome. J Orthop Surg Research 2013 Nov:8-40.

54. Clement ND, MacDonald D, Simpson A. The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. Knee Surgery. 2014.

55. Farrar JT, Berlin JA, Strom BL. Clinically important changes in acute pain outcome measures: a validation study. J Pain Symptom Manage. 2003 May;25(5):406–11.

56. Farrar JT, Portenoy RK, Berlin JA, Kinman JL, Strom BL. Defining the clinically important difference in pain outcome measures. Pain. 2000 Dec 1;88(3):287–94.

57. Tashjian RZ, Deloach J, Porucznik CA, Powell AP. Minimal clinically important differences (MCID) and patient acceptable symptomatic state (PASS) for visual analog scales (VAS) measuring pain in patients treated for rotator cuff disease. Journal of Shoulder and Elbow Surgery. 2009 Nov;18(6):927–32.

58. Reilly P, Macleod I, Macfarlane R, Windley J, Emery RJ. Dead men and radiologists don't lie: a review of cadaveric and radiological studies of rotator cuff tear prevalence. Ann R Coll Surg Engl. 2006 ed. 2006 Mar;88(2):116–21.

59. Fehringer EV, Sun J, VanOeveren LS, Keller BK, Matsen FA III. Full-thickness rotator cuff tear prevalence and correlation with function and co-morbidities in patients sixty-five years and older. Journal of Shoulder and Elbow Surgery. 2008 Nov;17(6):881–5.

60. Tempelhof S, Rupp S, Seil R. Age-related prevalence of rotator cuff tears in asymptomatic shoulders. Journal of Shoulder and Elbow Surgery. 1999 Jul;8(4):296–9.

61. Milgrom C, Schaffler M, Gilbert S, van Holsbeeck M. Rotator-cuff changes in

asymptomatic adults. The effect of age, hand dominance and gender. J Bone Joint Surg Br. 1995 Mar;77(2):296–8.

62. Moosmayer S, Smith HJ, Tariq R, Larmo A. Prevalence and characteristics of asymptomatic tears of the rotator cuff: AN ULTRASONOGRAPHIC AND CLINICAL STUDY. J Bone Joint Surg Br. 2009 Feb 1;91-B(2):196–200.

63. Baumgartner RN, Koehler KM, Gallagher D, Romero L, Heymsfield SB, Ross RR, et al. Epidemiology of sarcopenia among the elderly in New Mexico. Am J Epidemiol. 1998 ed. 1998 Apr 15;147(8):755–63.

64. Longo UG, Franceschi F, Ruzzini L, Rabitti C, Morini S, Maffulli N, et al. Histopathology of the Supraspinatus Tendon in Rotator Cuff Tears. The American Journal of Sports Medicine. 2007 Nov 30;36(3):533–8.

65. Schibany N, Zehetgruber H, Kainberger F. Rotator cuff tears in asymptomatic individuals: a clinical and ultrasonographic screening study. Europ J Radiology 2004 May 51: 263-268.

66. Yamaguchi K, Ditsios K, Middleton WD, Hildebolt CF, Galatz LM, Teefey SA. The demographic and morphological features of rotator cuff disease. A comparison of asymptomatic and symptomatic shoulders. J Bone Joint Surg Am. 2006 ed. 2006 Aug;88(8):1699–704.

67. Mall NA, Kim HM, Keener JD, Steger-May K, Teefey SA, Middleton WD, et al. J Bone Joint Surg Am. The American Orthopedic Association; 2010 Nov 17;92(16):2623–33.

68. Needell SD, Zlatkin MB, Sher JS, Murphy BJ, Uribe JW. MR imaging of the rotator cuff: peritendinous and bone abnormalities in an asymptomatic population. AJR American journal of roentgenology. American Public Health Association; 1996.

69. Chandnani V, Ho C, Gerharter J, Neumann C, Kursunoglu-Brahme S, Sartoris DJ, et al. MR findings in asymptomatic shoulders: a blind analysis using symptomatic shoulders as controls. Clinical Imaging. 1992 Jan;16(1):25–30.

70. Sher JS, Uribe JW, Posada A, Murphy BJ, Zlatkin MB. Abnormal findings on magnetic resonance images of asymptomatic shoulders. J Bone Joint Surg Am. The American Orthopedic Association; 1995 Jan 1;77(1):10–5.

71. Yamaguchi K, Tetro AM, Blam O, Evanoff BA, Teefey SA, Middleton WD. Natural history of asymptomatic rotator cuff tears: a longitudinal analysis of asymptomatic tears detected sonographically. J Shoulder Elbow Surg. 2001st ed. 2001 May-Jun;10(3):199–203.

72. Yamamoto A, Takagishi K, Osawa T, Yanagawa T, Nakajima D, Shitara H, et al. Prevalence and risk factors of a rotator cuff tear in the general population. J Shoulder Elbow Surg. 2009 ed. 19(1):116–20.

73. Minagawa H, Yamamoto N, Abe H, Fukuda M, Seki N, Kikuchi K, et al. Prevalence of symptomatic and asymptomatic rotator cuff tears in the general population: From mass-screening in one village. Journal of Orthopaedics. 2013 Mar;10(1):8–12.

74. Hinsley H, Nicholls A, Daines M, Wallace G, Arden N, Carr A. Classification of rotator cuff tendinopathy using high definition ultrasound. Muscles Ligaments Tendons J. 2014 Jul;4(3):391–7.

75. Moosmayer S, Tariq R, Stiris M, Smith H-J. The Natural History of Asymptomatic Rotator Cuff Tears. J Bone Joint Surg Am. 2013 Jul 17;95(14):1249.

76. Yamamoto A, Takagishi K, Kobayashi T, Shitara H, Osawa T. Factors involved in the presence of symptoms associated with rotator cuff tears: a comparison of asymptomatic and symptomatic rotator cuff tears in the general population. Journal of Shoulder and Elbow Surgery. 2011 Oct;20(7):1133–7.

77. Harris JD, Pedroza A, Jones GL, MOON (Multicenter Orthopedic Outcomes Network) Shoulder Group. Predictors of pain and function in patients with symptomatic, atraumatic full-thickness rotator cuff tears: a time-zero analysis of a prospective patient cohort enrolled in a structured physical therapy program. The American Journal of Sports Medicine. 2012 Feb;40(2):359–66.

78. Constant CR, Murley AG. A Clinical Method of Functional Assessment of the Shoulder. Clinical Orthopaedics and Related Research. 1987 Jan 1;214:160.

79. Constant CR, Gerber C, Emery RJH, Søjbjerg JO, Gohlke F, Boileau P. A review of the Constant score: Modifications and guidelines for its use. Journal of Shoulder and Elbow Surgery. 2008 Mar;17(2):355–61.

80. Bankes MJK, Crossman JE, Emery RJH. A standard method of shoulder strength measurement for the Constant score with a spring balance. Journal of Shoulder and Elbow Surgery. 1998 Mar;7(2):116–21.

81. Walton MJ, Walton JC, Honorez LA, Harding VF, Wallace WA. A comparison of methods for shoulder strength assessment and analysis of Constant score change in patients aged over fifty years in the United Kingdom. J Shoulder Elbow Surg. 2007 ed. 2007 May-Jun;16(3):285–9.

82. Thomas M, Dieball O, Busse M. [Normal values of the shoulder strength in dependency on age and gender--comparison with the constant, UCLA, ASES scores and SF36 health survey]. Z Orthop Ihre Grenzgeb. 2003 Mar;141(2):160–70.

83. Kim HM, Teefey SA, Zelig A, Galatz LM, Keener JD, Yamaguchi K. Shoulder strength in asymptomatic individuals with intact compared with torn rotator cuffs. J Bone Joint Surg Am. 2009 Feb;91(2):289–96.

84. Keener JD, Steger-May K, Stobbs G, Yamaguchi K. Asymptomatic rotator cuff tears: Patient demographics and baseline shoulder function. Journal of Shoulder and Elbow Surgery. 2010 Dec;19(8):1191–8.

85. Nakajima D, Yamamoto A, Kobayashi T, Osawa T, Shitara H, Ichinose T, et al. The effects of rotator cuff tears, including shoulders without pain, on activities of daily living in the general population. J Orthop Sci. 2012 Mar;17(2):136–40.

86. Angst F, Schwyzer HK, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function. Arthritis care & Research 2011 Nov;63(11):174-188.

87. Briggs T (2012). Getting it right first time. Improving the quality of orthopaedic care within the National Health Service in England.

88. Judge A, Murphy RJ, Maxwell R, Arden NK, Carr AJ. Temporal trends and geographical variation in the use of subacromial decompression and rotator cuff repair of the shoulder in England. The Bone & Joint Journal. 2014 Jan;96-B(1):70–4.