PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prevalence of rotator cuff tendon tears and symptoms in a Chingford general population cohort, and the resultant impact on United Kingdom health services: A cross-sectional observational study.
AUTHORS	Hinsley, Hannah; Ganderton, Charlotte; Arden, Nigel; Carr, Andrew

VERSION 1 – REVIEW

REVIEWER	Smythe, A Monash University, Physiotherapy
REVIEW RETURNED	04-Dec-2021

GENERAL COMMENTS	Overall, this was a very interesting, important and well written study addressing important clinical questions. I think that with some revision it will be a great publication! 1.) Somewhere in the introduction or methods can you please define terminology of rotator cuff tendinopathy. Terminology around this diagnoses in contentious and arguments can be made that it's interchangeable with rotator cuff related pain, subacromial impingement syndrome etc. You define the imaging criteria for diagnosis well but don't explain what rotator cuff tendinopathy is in terms of your study. 2.) Can you please elaborate on why you used a cut off of 47/48 on the OSS to classify a shoulder as symptomatic. You partially addressed this in limitations but I think that it is the basis of a good portion of your results/findings you need some evidence or reasoning of why you chose this cut off. I say this because there is a strong argument that asymptomatic people quite often score less than 47 and that with increasing age this 'normal' decreases (see ref below). Also your reference 35 states; "It should, however, be noted that with all outcome scores, scores tend to worsen with age [36, 37]. Therefore, for each condition or type of surgery that is studied, in elderly patients, a "normal" score may be somewhat less than 48." If you don't address this it is a very big limitation and weakens the study. Clement, Nicholas, & Court-Brown, Charles. (2014). Oxford shoulder score in a normal population. International Journal of Shoulder Surgery, 8(1), 10–14. https://doi.org/10.4103/0973-6042.131849. 3.) Were patients asked to complete an OSS for each shoulder or one to cover both shoulders, and how was this addressed statistically? 4.) References – please choose to go with full journal names or abbreviations but don't do both 5.) The remainder are small changes Line 103 "Severity of symptoms it not related to the severity of the

pathology" should be severity of symptoms was not related to the ….

Line 105 – Don't start a sentence with number symbol, consider rewording

Line 133 – remove comma after however

Line 141 – 'a risk' instead of 'at risk'

Line 145-147 – please elaborate on what you mean by the individual having an influence on symptoms. It is a bit vague.

Figure 2 is missing

Line 287 – Don't start sentence with number symbol

Line 309 – 'decline in tendon tissue' is a bit vague. This could talk to reducing resilience, reducing volume, general degradation. Please change wording to be a bit more specific.

Line 311 – remove 'functional tasks' and stay with higher cumulative loading.

Line 386-388 – again please elaborate on what you mean when you say ' the effect the individual has on symptom presentation'. Is this more speaking to demographics of the individual or just non-imaging factors?

Line 414-415 – please state somewhere this is in a female only population as it can be a bit misleading conclusion. Something like "affecting 22.1% of women over the age of 60"

Line 422 "non-torn tendon"

Ref 1 – third line please fix syntax

Ref 10 – journal name has both abbreviation and full name, please choose one

REVIEWER	Ingwersen, Kim Gordon Sygehus Lillebalt Vejle Sygehus, Department of Physio- and Occupational Therapy
REVIEW RETURNED	17-Dec-2021

GENERAL COMMENTS	First, I would like to acknowledge the relevance and important aspects of this manuscript. This trial gives an example of how common and how heavily rotator cuff tears impact the health care system. Furtheremore, the challenge diagnosing shoulder discomfort is highlighted in this manuscript with the verification that almost 50% of full-thickness rotator cuff tears are asymptomatic. However, I have some questions and comments. 1. The objectives in the abstract and in the introductions aren´t entirely balanced. 2. The use of rotator cuff tendinopathy in the title and tears in the objectives are confusing. 3. The majority of the trial are concerning rotator cuff full-thickness tears, and only partly the abnormal/partial tear groups. It is well established that abnormal tendon structures are difficult to define on ultrasound, and even though the authors present the reliability between their raters, the difficulties in definition of a normal or abnormal tendon is only minimally discussed. Please elaborate on this. 4. The trial is based upon the Chingford Study. The recruitment to this specific trial is only minimally described, and with only 463 out of the original 1003 cohort participants, are more detailed explanation of what was done to ensure full participation would be preferred, in order to be able to repeat this trial. Furthermore, only a minimum of information is given about the Chingford Study, and no demographic description is given about the included woman, compared to the woman not included – it is therefore difficult to know whether the cohort are representative of the UK general population.

2

	5. The cut-off value the authors are using for the dichotomized OSS could be questioned. Some of the items in OSS could be scored "not perfect" due to other problems then shoulder pain. The authors state that a 3 point change was tested – As I understand this, they tested if the cut-off score of <45 made a difference. Please confirm this 6. In relation to the study limitations, the argument that only woman is included in this cohort, but this will not bias the results, is only minimally supported by two previous studies. I would prefer a more detailed consideration upon this limitation. Furthermore, the authors state that no known association exist between shoulder pain and other medical co-morbidities without being able to put an reference to this statement. Several trials have shown associations between diabetes, obesity and metabolic syndrome (related to medical co-morbidities) and musculoskeletal pain. Taking this into consideration, there could possibly be a survival bias in the cohort. 7. In line 341 you state that individuals participating in the medical examination was selected at random – please elaborate on how you did this, og what supports that the participating women was not different from the woman who choose not to participate in the medical examination. Stating that BMI and age is the same does not mean that their wasn´t other factors that could cause a difference. 8. The authors properly states that there is a risk of overreporting pathology, as they were aware of the OSS results. I would like the authors to elaborate upon this, as this especially among the groups Normal versus abnormal/partiel tears can have an large influence upon the which group the examiner defines. 9. In line 354-355 the authors state that the removal of asymptomatic shoulders, would reduce the backgraound noise from other potential painful conditions. I am not sure what the authors means by this – please elaborate.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. A Smythe, Monash University, Lively Physiotherapy

Comments to the Author:

Comment 1: Overall, this was a very interesting, important and well written study addressing important clinical questions. I think that with some revision it will be a great publication!

Comment: Somewhere in the introduction or methods can you please define terminology of rotator cuff tendinopathy. Terminology around this diagnoses in contentious and arguments can be made that it's interchangeable with rotator cuff related pain, subacromial impingement syndrome etc. You define the imaging criteria for diagnosis well but don't explain what rotator cuff tendinopathy is in terms of your study.

Response 1: Additional information has been provided to explain the appearance of each tendon classification based on the prior research and publication by Hinsley (2014). A figure legend has been added to Figure 1 to improve clarity. Furthermore, we have removed the word 'tendinopathy' unless we are referring to abnormal tendon (pathological) in the presence of symptoms and referred to tendon classifications as 'normal' or 'abnormal' being tendon pathology of varying severity (e.g. full thickness tear).

Ref: Hinsley H, Nicholls A, Daines M, et al. Classification of rotator cuff tendinopathy using high definition ultrasound. Muscles Ligaments Tendons Journal 2014;4(3):391-7. [published Online First: 2014/12/10]

Comment 2: Can you please elaborate on why you used a cut off of 47/48 on the OSS to classify a shoulder as symptomatic. You partially addressed this in limitations but I think that it is the basis of a good portion of your results/findings you need some evidence or reasoning of why you chose this cut off. I say this because there is a strong argument that asymptomatic people quite often score less than 47 and that with increasing age this 'normal' decreases (see ref below). Also your reference 35 states; "It should, however, be noted that with all outcome scores, scores tend to worsen with age [36, 37]. Therefore, for each condition or type of surgery that is studied, in elderly patients, a "normal" score may be somewhat less than 48."

If you don't address this it is a very big limitation and weakens the study.

Clement, Nicholas, & Court-Brown, Charles. (2014). Oxford shoulder score in a normal population. International Journal of Shoulder Surgery, 8(1), 10–14.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4049034/pdf/IJSS-8-10.pdf

Response 2: The Oxford Shoulder Score was originally designed to look at change between pre and post shoulder surgery. In this study, we did not used it for this purpose. Instead, we use it to detect any individual who was unable to perform an activity to the full, or who has pain at any given time. We have added the following information into the manuscript:

"Symptoms were defined using the Oxford shoulder score34 35. This was chosen for what the authors believed represented the best content and construct validity as applicable to the study as it covers a range of symptoms (both relating to pain and function) over a 4-week time period, and also allows discriminate ability. Binary symptoms were defined by dichotomising the Oxford shoulder score34 35 where, any non-perfect score (≤ 47/48) was classified as symptomatic. The cut off at 47 was used to determine symptoms as we were not looking for significant changes, rather, the ability to detect any individual who was unable to perform an activity to the full, or who has pain at any given time. This showed good correlation with binary pain questions and the NRS and was not statistically different to the results using a 3-point gap." Lines 218-226. It should also be noted that the authors have run an analysis on results with a cut off of a three-point difference (45/48) and a one-point difference (47/48), and results were identical.

Comment 3: Were patients asked to complete an OSS for each shoulder or one to cover both shoulders, and how was this addressed statistically?
Response 3: The OSS was completed for each shoulder. This has been clarified in the method: "…shoulder ultrasound examination was performed on both shoulders (left and right) in 463 women…" (lines 172-173)

Comment 4: References – please choose to go with full journal names or abbreviations but don't do both
Response 4: Thank you. The full journal name has been provided for all references.

Comment 5: Line 103 "Severity of symptoms it not related to the severity of the pathology" should be severity of symptoms was not related to the ….
Response 5: Thank you. This has been amended to read "severity of symptoms was not related" (line 105)

Comment 6: Line 105 – Don't start a sentence with number symbol, consider rewording
Response 6: This has been amended: "In the cohort 8.9% had seen their GP with shoulder pain and a full-thickness rotator cuff tear, 18.8% with an abnormality and 29.3% overall" (line 108-109)

Comment 7: Line 133 – remove comma after however
Response 7: Correction has been made.

Comment 8: Line 141 – 'a risk' instead of 'at risk'
Response 8: This has been amended.

Comment 9: Line 145-147 – please elaborate on what you mean by the individual having an influence on symptoms. It is a bit vague.

Response 9: This section has been elaborated on and the following information included "However, all studies investigating symptom association have looked at isolated shoulders, and have not considered that the individual, has two shoulders. It is therefore plausible that there may be the presence of other physical or psychological factors unique to the individual rather than the specific shoulder that may have an influence on symptom presentation, rather than solely the underlying pathology." (lines 147-151)

Comment 10: Figure 2 is missing

Response 10: Apologies for this. Figure 2 has been uploaded with the revised manuscript submission.

Comment 11: Line 287 – Don't start sentence with number symbol

Response 11: This has been amended to read "In this cohort, 8.9% (41/463)…" (line 310)

Comment 12: Line 309 – 'decline in tendon tissue' is a bit vague. This could talk to reducing resilience, reducing volume, general degradation. Please change wording to be a bit more specific.

Response 12: The authors agree that this could be confusing so the wording has been altered and a reference cited: "Prevalence was found to increase with every decile of age, and the relative risk of having a full thickness tear increased more than two-fold between the 65-69 and >80 age groups, suggesting age related change18" (line 331).

Comment 13: Line 311 – remove 'functional tasks' and stay with higher cumulative loading.

Response 13: This has been amended according to your suggestion.

Comment 14: Line 386-388 – again please elaborate on what you mean when you say ' the effect the individual has on symptom presentation'. Is this more speaking to demographics of the individual or just non-imaging factors?

Response 14: Reference to the 'individual' has been clarified: "This is the first study that has looked at individuals as entities, rather than shoulders, and has highlighted the effect the individual has on symptom presentation, which could include physical and psychological factors unique to that individual – not solely the presence of tendon pathology on imaging. It is also the first study to look at the impact on health services" (lines 415-418).

Comment 15: Line 414-415 – please state somewhere this is in a female only population as it can be a bit misleading conclusion. Something like "affecting 22.1% of women over the age of 60"

Response 15: Thank you for improving the clarity of this conclusion. We have made the amendments as per your recommendations.

Comment 16: Line 422 "non-torn tendon"

Response 16: Thank you, this has been changed.

Comment 17: Ref 1 – third line please fix syntax. Ref 10 – journal name has both abbreviation and full name, please choose one

Response 17: Thank you for identifying these reference errors. They have been corrected and all references have the full journal name as per reviewer 2.

Reviewer: 2
Dr. Kim Gordon Ingwersen, Sygehus Lillebalt Vejle Sygehus

Comments to the Author:

Dear authors,

Comment 18: First, I would like to acknowledge the relevance and important aspects of this manuscript. This trial gives an example of how common and how heavily rotator cuff tears impact the health care system. Furthermore, the challenge diagnosing shoulder discomfort is highlighted in this manuscript with the verification that almost 50% of full-thickness rotator cuff tears are asymptomatic. However, I have some questions and comments.

Response 18: Thank you. We hope that we have been able to address your queries below.

Comment 19: The objectives in the abstract and in the introductions aren´t entirely balanced.

Response 19: Agreed. The abstract objectives have been changed to more closely reflect those in the introduction: "To define the population prevalence of rotator cuff tears and test their association with pain and function loss; determine if severity symptom correlates with tear stage severity, and quantify the impact of symptomatic rotator cuff tears on primary health care services, in a general population cohort of women." (lines 81-84)

Comment 20: The use of rotator cuff tendinopathy in the title and tears in the objectives are confusing.

Response 20: The title has been changed and the word 'tendinopathy' has been changed to 'tendon tears and symptoms'. The new title is Prevalence of rotator cuff tendon tears and symptoms in a Chingford general population cohort, and the resultant impact on United Kingdom health services: A cross-sectional observational study.

Comment 21: The majority of the trial are concerning rotator cuff full-thickness tears, and only partly the abnormal/partial tear groups. It is well established that abnormal tendon structures are difficult to define on ultrasound, and even though the authors present the reliability between their raters, the difficulties in definition of a normal or abnormal tendon is only minimally discussed. Please elaborate on this.

Response 21: We have added the following information regarding the agreement between raters of a small intra-observer study completed: "To overcome this, a small intra-observer study was performed on 18 participants who were re-scanned at a different time-period. The examiner was blinded to all pervious results and shoulder scores. Overall agreement gave a weighted kappa score of 0.915 (p<0.001)". Line 374-377. We have also provided addition information on the appearance of tendon in each classification based on prior research of Hinsley et al. (2014). This paper recognises the reduced sensitivities of detected partial thickness tear, but proposes grouping together all abnormailies short of full-thickness tears is a reasonable valid classification – see Figure 1 Legend that states: "Tendon classification on ultrasound: (i) normal tendon: normal homogenous appearance throughout with no abnormality at the enthesis; (ii) abnormal tendon: loss of homogenous appearance and abnormal ragged enthesis +/- enlarged fluid-filled bursa or partial thickness tear; (iii) full thickness tear (0-2.5cm): lucent patch through the full thickness of the tendon with tear size defined as its width in the sagittal plane (iv) full-thickness tears (>2.5cm): Evidence of large defect or no evidence of tendon tissue present.

Ref: Hinsley H, Nicholls A, Daines M, et al. Classification of rotator cuff tendinopathy using high definition ultrasound. Muscles Ligaments Tendons Journal 2014;4(3):391-7. [published Online First: 2014/12/10]

Comment 22: The trial is based upon the Chingford Study. The recruitment to this specific trial is only minimally described, and with only 463 out of the original 1003 cohort participants, are more detailed explanation of what was done to ensure full participation would be preferred, in order to be able to repeat this trial. Furthermore, only a minimum of information is given about the Chingford Study, and no demographic description is given about the included woman, compared to the woman not included – it is therefore difficult to know whether the cohort are representative of the UK general population.

Response 22: Additional information has been added to the 'design' section of the manuscript. The authors feel that the information provided is comprehensive, and appropriate references given to prior research (lines 165-177):
"Participants in this cross-sectional observational study were involved in the larger Chingford 1000 women study. This is an ethically approved well described prospective population-based longitudinal study of osteoarthritis and osteoporosis comprising 1003 white Caucasian women, derived from the register of a large general practice in Chingford, North London31-33. The cohort was recruited in 1989 where the women were aged 44-67. They have been characterised as representative of women in the UK general population with respect to weight, height, and smoking characteristics. The cohort has been subsequently listed by the National Institute for Health Research as an important epidemiological recourse. This study took place at the Chingford 20 year follow up visit where 516 of the original 1003 cohort attended (158 women had died, 111 were unable to attend, 218 had moved away or been lost to follow up). A musculoskeletal assessment, including the Oxford shoulder score, and shoulder ultrasound examination was performed on both shoulders (left and right) in 463 women (Out of the 515, 52 attended but did not have a shoulder assessment due to lack of assessor, and 1 did not complete an Oxford shoulder score)." (Lines 163-174)

Comment 23: The cut-off value the authors are using for the dichotomized OSS could be questioned. Some of the items in OSS could be scored "not perfect" due to other problems then shoulder pain. The authors state that a 3-point change was tested – As I understand this, they tested if the cut-off score of <45 made a difference. Please confirm this
Response 23: Thank you, this has been addressed as per reviewer 1 Comment 2/Response 2.

Comment 24: In relation to the study limitations, the argument that only woman is included in this cohort, but this will not bias the results, is only minimally supported by two previous studies. I would prefer a more detailed consideration upon this limitation. Furthermore, the authors state that no known association exist between shoulder pain and other medical co-morbidities without being able to put an reference to this statement. Several trials have shown associations between diabetes, obesity and metabolic syndrome (related to medical co-morbidities) and musculoskeletal pain. Taking this into consideration, there could possibly be a survival bias in the cohort.
Response 24: The authors are not aware of any literature that suggests that rotator cuff tears are more likely in females than males, and therefore it is possible to generalise the results of this study to the general population (inclusive of males). In regard to survival bias, if we were to lose the more 'frail' people to the 20-year follow up, it is theoretically possible that the ones left may have less pathology, and therefore prevalence may be underestimated, however there are no studies in the literature that suggest rotator cuff tears are associated with general medical conditions. Furthermore, the cohort was originally investigated with the primary focus of osteoporosis, and not shoulder symptoms, thus any continued participation is not driven by shoulder symptoms. The following has been added to the manuscript: "Potential survival bias is introduced by the cohort being in its 20th year. If a greater proportion of individuals with pathology were lost to follow up this may cause us to under-estimate any association, however, no known associations exist in the literature between rotator cuff tears and other medical co-morbidities." (lines 357-360).

Comment 25: In line 341 you state that individuals participating in the medical examination was selected at random – please elaborate on how you did this, og what supports that the participating women was not different from the woman who choose not to participate in the medical examination. Stating that BMI and age is the same does not mean that their wasn´t other factors that could cause a difference.
Response 25: A correction has been made to this. The word 'selected' has been removed and additional information on why not all individuals had the 20-year shoulder examination was provided: "Furthermore, only 463/516 individuals that attended the year-20 study underwent a shoulder examination due to lack of an examiner being present at these follow up appointments". (lines 362-

364)

Comment 26: The authors properly states that there is a risk of overreporting pathology, as they were aware of the OSS results. I would like the authors to elaborate upon this, as this especially among the groups Normal versus abnormal/partial tears can have a large influence upon the which group the examiner defines.

Response 26: Additional information has been provided regarding the risk of overreporting pathology: "To overcome this, a small intra-observer study was completed, and an additional ultrasounds scan was performed on 18 willing participants. The examiner was blind to all pervious results and shoulder scores. Overall agreement gave a weighted kappa score of 0.915 (p<0.001)." (lines 374-377)

Comment 27: In line 354-355 the authors state that the removal of asymptomatic shoulders, would reduce the background noise from other potential painful conditions. I am not sure what the authors means by this – please elaborate.

Response 27: This section has been revised to improve clarity for readers: "The effect of tear size on symptom severity may have been underestimated in this study. The inability to transform the complete data set due to the skew of the OSS data, meant all asymptomatic shoulders had to be removed. Pain severity in the presence of a tear was then compared to a pain severity in a normal (no tendon pathology) shoulder. We recognise that there may be many causes of shoulder pain (e.g., rheumatological causes) and therefore referencing against all causes of painful shoulder may represent the contribution of rotator cuff tear to the symptoms." (lines 379-384)

VERSION 2 – REVIEW

REVIEWER	Smythe, A Monash University, Physiotherapy
REVIEW RETURNED	26-Mar-2022

| GENERAL COMMENTS | Good job all. This has improved since previous version. My main issue though is still the dichotomisation of OSS as perfect vs non-perfect.

Line 108,109: Sentence is confusing. Is it saying 29.3% had seen their GP overall?

Line 225: OSS showed good correlation with binary pain questions – can you please show this in results.

LLine 288: Table 3. Title is misleading. It should be titled Ultrasound findings demographics or similar as in line 283 you state: "For the 289 symptomatic shoulders the full OSS was reported (Table 3)". I think it would be of benefit to see the table you have provided but also add in OSS scores, eg. mean for each US group

It would be beneficial to extend table 4 to include distribution of shoulder symptoms against FFT (as you have done), abnormal tendons, and normal tendons.

Line 350: "The major strength of this study is that it uses a large general population cohort, and therefore not subject to selection bias". I would disagree, as it is women only, I don't think you can report it as a general population if half of the general population is not included. |

	Line 390: You really need to convince me that perfect vs non-perfect scores is valid as your references stated in my previous rounds of comments state that scores for normal shoulders drop with increasing age. If you ran a comparison with 3-point change I want to see the data in your study to convince me that 3-point change didn't change outcomes. The other thing that would really help me accept this dichotomisation is if you had: 1) a table or maybe even better a graph depicting OSS scores mean/Cis against age groups, BMI etc. Maybe also splitting into symptomatic and asymptomatic or; 2) table two add symptomatic vs asymptomatic into table.
	Line 446: I would suggest not labelling this a general population study if it is women only
	Overall there are only a few minor changes but I want you to show me more data so that any weakness/pattern caused by dichotomising results of the OSS is clear and can be easily interpreted by the reader.

REVIEWER	Ingwersen, Kim Gordon Sygehus Lillebalt Vejle Sygehus, Department of Physio- and Occupational Therapy
REVIEW RETURNED	24-Mar-2022

GENERAL COMMENTS	Dr Hannah Hinsley, Professor Nigel Arden, Dr Charlotte Ganderton and Professor Andrew Carr. I would thank you for your comments and amendments to my previous review. I have no further suggestions or comments.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Dr. Kim Gordon Ingwersen, Sygehus Lillebalt Vejle Sygehus

Comments to the Author 1:
Dear Dr Hannah Hinsley, Professor Nigel Arden, Dr Charlotte Ganderton and Professor Andrew Carr.
I would thank you for your comments and amendments to my previous review. I have no further suggestions or comments.
Response 1:
Thankyou for your kind review.

Reviewer: 1
Dr. A Smythe, Monash University, Lively Physiotherapy

Comments to the Author 1:
Good job all. This has improved since previous version. My main issue though is still the dichotomisation of OSS as perfect vs non-perfect.
Response 1:
Thank you for your kind and considered review. We have addressed all concerns and have added additional information regarding dichotomisation of OSS as perfect and non-perfect scores. Please see Response 3 and 7 for information regarding this.

Comment 2: Line 108,109: Sentence is confusing. Is it saying 29.3% had seen their GP overall?
Response 2:

The following information has been added to improve clarity: "In the cohort 8.9% had seen their GP with shoulder pain and a full-thickness rotator cuff tear, 18.8% with shoulder pain and an abnormality and 29.3% with shoulder pain." (lines 108-109)

Comment 3: Line 225: OSS showed good correlation with binary pain questions – can you please show this in results.
Response 3:
Thank you. We have included the following additional information: "We validated this by running a Pearson correlation sub analysis between the OSS pain subset with the NRS (R=0.816, p<0.001, 95% CI 0.793-0.836) and a simple binary question (R=0.812, p<0.001, 95% CI 0.789-0.833), and the full OSS with a binary pain question (R=0.759, p<0.001, 95% CI 0.730-0.785). Furthermore, we re-ran the analysis using a 3-point difference to reflect a clinically significant difference between groups and the results were not significantly different." (lines 224-229)

Comment 4: Line 288: Table 3. Title is misleading. It should be titled Ultrasound findings demographics or similar as in line 283 you state: "For the 289 symptomatic shoulders the full OSS was reported (Table 3)". I think it would be of benefit to see the table you have provided but also add in OSS scores, eg. mean for each US group
Response 4:
Thank you for your ideas here. We have updated the title of the table "Table 3. Demographics of the 289 symptomatic shoulders".

Comment 5: It would be beneficial to extend table 4 to include distribution of shoulder symptoms against FFT (as you have done), abnormal tendons, and normal tendons.
Response 5:
Thank you for your suggestion. We have included the following information within Table 4. Please see a copy of the updated below and lines 312-314 of the manuscript:

Table 4. Distribution of individual shoulder symptoms according to the presence of full-thickness tears or tendon abnormalities

	No Symptoms	Unilateral Symptoms	Bilateral Symptoms	Total
Bilateral No FTT	226	71	63	360
Unilateral FTT	33	25	24	82
Bilateral FTT	10	3	8	21
Bilateral normal	131	28	28	187
Unilateral abnormality	72	34	28	134
Bilateral abnormality	66	37	39	142
Total	269	99	95	463

Comment 6: Line 350: "The major strength of this study is that it uses a large general population cohort, and therefore not subject to selection bias". I would disagree, as it is women only, I don't think you can report it as a general population if half of the general population is not included.
Response 6:
Thank you. The word 'general' has been removed here: "The major strength of this study is that it uses a large population-based cohort, and therefore not subject to selection bias." (lines 357-358)

Comment 7: Line 390: You really need to convince me that perfect vs non-perfect scores is valid as your references stated in my previous rounds of comments state that scores for normal shoulders drop with increasing age. If you ran a comparison with 3-point change I want to see the data in your study to convince me that 3-point change didn't change outcomes. The other thing that would really help me accept this dichotomisation is if you had: 1) a table or maybe even better a graph depicting OSS scores mean/Cis against age groups, BMI etc. Maybe also splitting into symptomatic and asymptomatic or; 2) table two add symptomatic vs asymptomatic into table.
Response 7:

1. "We validated this by running a Pearson correlation sub analysis between the OSS pain subset with the NRS (R=0.816, p<0.001, 95% CI 0.793-0.836) and a simple binary question (R=0.812, p<0.001, 95% CI 0.789-0.833), and the full OSS with a binary pain question (R=0.759, p<0.001, 95% CI 0.730-0.785). Furthermore, we re-ran the analysis using a 3-point difference to reflect a clinically significant difference between groups and the results were not significantly different." (lines 224-229)

2. When the same analysis was performed using a 3-point change in the OSS to define symptoms the results were not statistically different and compared to normal tendons were as follows: Abnormal/Partial tears 1.793 (OR 1.936, 95%CI 1.374-2.726); full-thickness tears 0-2.5cm 2.098 (OR 2.506, 95%CI 1.513-4.150); and full-thickness tears >2.5cm 3.924 (OR 9.678, 95%CI 4.784-19.580).

3. With regards to changes in normal OSS with age – the authors cannot agree with the paper refenced to. The authors have used a biased MSK cohort. Although they have shown differing median scores with age – between 60-80 there is no difference represented by overlapping CIs. This would support our use in this cohort. Furthermore, they have proceeded to statistically analyse mean scores which is not valid in highly skewed data and thus they cannot demonstrate statistical differences.

This is an area that we spent a long time looking at how to analyse the data. Our decision was based on the support of our statistician who was under the supervision of Professor of Translational Statistics, Andrew Judge, at the time.

Comment 8: Line 446: I would suggest not labelling this a general population study if it is women only
Response 8:
Reference to 'general' has been removed. "In conclusion, this population-based study has demonstrated…"(lines 453)

Comment 9: Overall there are only a few minor changes but I want you to show me more data so that any weakness/pattern caused by dichotomising results of the OSS is clear and can be easily interpreted by the reader.
Response 9:
Thank you for your thorough review of this manuscript. As authors, we hope that we have been able to address your concerns sufficiently and improve our submission to a status worthy of publication.

VERSION 3 – REVIEW

REVIEWER	Smythe, A
	Monash University, Physiotherapy
REVIEW RETURNED	05-Jun-2022

GENERAL COMMENTS	Thankyou for addressing my points in the last revision. I think it makes for a very good piece of research.