Author’s response to reviews

Title: Reliability and methodology of quantitative assessment of harvested and unharvested patellar tendons of ACL injured athletes using ultrasound tissue characterization

Authors:

Carla Pereira (carla.pereira@aspetar.com; carlasonsino@hotmail.com)

Rafael Santos (rafael.santos@aspetar.com)

Rodney Whiteley (rodney.whiteley@aspetar.com)

Taija Finni (taija.finni@jyu.fi)

Version: 1 Date: 18 Feb 2019

Author’s response to reviews:

Alex Scott (Reviewer 1): This paper examines the reliability of UTC values from the patellar tendon in thirty-seven male professional athletes, all of whom underwent different forms of ACL repair.

Nineteen participants had their harvested and unharvested patellar tendons scanned twice by the same examiner.

Twenty-seven participants had their patellar tendons scanned twice by two different examiners.

COMMENT 1: Recommend the Strobe checklist for observational studies be followed. For example, inclusion/exclusion criteria are missing, and it's not immediately clear that all 37 patients underwent an ACL repair until one gets to the results.

REPLY: We used the Strobe checklist for cross sectional study. The reason for that is because the study is conducted in a clinical setting, thus all the ACL patients who underwent assessment in the ACL Assessment center of Aspetar and signed the consent form could had been included in this sample, so there were no exclusion criteria. Additionally, because we haven’t performed any comparison within patient, even the data of patients with bilateral ACL injury/repair could be included.

ACTION: We have added in the 1st paragraph of methods that they were ACL patients and that the same patient might have undergone repetitive assessments (page 4).
COMMENT 2: It's also a bit difficult to understand the number of measures done on each participant - there must be some overlap between the 19 and the 27 in terms of their participation in the inter- and intra-rater reliability, could this be indicated in a table or Venn diagram?

REPLY: There is an overlap between patients and measurements. These patients were in treatment in Aspetar. They were assessed in the ACL Ax center every 6 weeks regardless of this study, and the period of data acquisition for this paper lasted around 6 months. Thus, we had data of the same patient at different stages, as for example before surgery and after harvesting the patellar tendon due to the ACL reconstruction. Also, to consider that both patellar tendons of the same subject could had been used for analysis (the ACL involved leg - and the non-ACL uninvolved leg-).

ACTION: We modified the way we mentioned the number of subjects and patellar tendons in the abstract (page 2), in the methods section (page 5), and in all the tables.

-We included a diagram (Figure 1)

COMMENT 3: For the patients who participated in both studies, which of the two measures from the intra-rater reliability study were used for the analysis of inter-rater reliability?

REPLY: Always the data of the first acquisition of the intra-rater reliability was selected to be included in the inter-rater reliability analysis.

ACTION: Explanation was included in methods (page 8). “In the occasion of duplicated analysis of the same patellar tendon, always the data from the first acquisition of the most experienced examiner was selected for inter-rater reliability analysis.”

COMMENT 4: Could the authors please review the length calculations on page 6? If each scan is 0.2mm in thickness (600 scans over 12 cm), then are these length calculations accurate, or were they derived differently.

REPLY: The UTC imaging acquires 600 sequential transversal images of 0.02cm thickness resulting in a total length of 12cm of data acquisition.

The length measurement we adopted in this study considers the distance between the patellar apex (1st area of interest) and the notch of the tibia (the more distal contour of each tendon) as the patellar tendon length. For instance, if someone presented 4.64 cm of tendon length it means that 232 images were present between patellar apex and notch.

ACTION: We changed the unit from mm to cm on the description of the UTC imaging capture (page 4).
We included in Fig 3. an example of the patellar tendon measurement (page 7).

COMMENT 5: I appreciate that there seem to be different formulas floating around the literature for the calculation of MDC, and this is likely something we need to agree on, at least within the UTC field. Could the editor send the paper for statistical review to decide if the method used here is appropriate? The MDC values are very small (less than 1%) which is counter-intuitive and different from what has previously been reported. Which ICC model was used? Could the discussion section compare the current findings to previous ones?

REPLY: According to the 3rd reviewer of this manuscript, Marijke Welvaert: “Upon request from the Editor, I focused my review on the statistical aspects of the manuscript. Overall, I think the statistics are well reported, the chosen method is appropriate, and the interpretation of the results support the presented numerical outcomes”.

Thank you so much for your remark. I went back to all the data and reviewed all the analysis. All the methods used to calculate these parameters are in agreement with the literature (Beckerman et al., 2001; de Jonge, Tol, & Weir, 2015; S. I. Docking & Cook, 2016; Sean I. Docking, Rosengarten, Daffy, & Cook, 2015; Rosengarten et al., 2014; Wezenbeek et al., 2017). However, I had missed two extremely important steps that changed completely the results of SEM and MDC based on the ICC. Our raw data for echo-types is extracted as decimals instead of percentage as presented in this paper. Therefore, the MDC results were so small, it was missing *100 in one step and a pair of brackets including the *100 in another step of the calculations). My sincere apologies, and again, my high appreciation for your comment.

Majority of papers calculating MDC were done in Achilles tendons, the one from Docking that calculated MDC in patellar tendons presented data for echo-types I + II, as aligned structure, and echo-types III + IV as disorganized structure. To the best of our knowledge, nobody has calculated MDC for harvested patellar tendons, or even unharvested patellar tendons but considering the echo-types separately.

Thus, we will not add comparisons of it in the discussion.

ACTION: We have reviewed every single calculation performed and found mistakes that now has been fixed.

MDC was now calculated exclusively with the data that were acquired in two different days, with one day interval, by a single examiner.
Additionally, we decided to present the MDC results as it has been previously presented by de Jonge et al. 2015, Docking et al. 2015 and Wezenbeek et al. 2017 to facilitate understanding.

Our modified MDC results = 7.5 % for echo-type I, 6.9 % for echo-type II, 4.8 % for echo-type III and 2 % for echo-type IV. For unharvested tendons, the MDC was 14.1 % for echo-type I, 10.6 % for echo-type II, 6.3 % for echo-type III and 1.2 % for echo-type IV.

COMMENT 6: The explanation for what UTC values supposedly map to, histologically, is not accepted by many within the community, mainly because ultrasound does not provide images at a histological scale/resolution. Also, the original validation study in horses was done with different parameters than those used here. Could the discussion be balanced somewhat?

REPLY: Thank you for your comment.

Even though some may not accept the use of UTC based on the histological comparisons performed in horses’ tendon, this method has been clinically validated and has been widely used for clinicians to explore tendon pathology. Differences in echo-types distribution in the presence of pathology for involved and uninvolved limbs, load responses, relation with clinical symptoms, and response to treatment are some examples of successful use of UTC.

ACTION: A note in caution chapter has been added to highlight the lack of a histologically validated study in human tendons.

Peter Preston, Masters (Reviewer 2): BMC Sports Science, Medicine and Rehabilitation

Clinimetrics and normative values of quantitative assessment of harvested patellar tendons of athletes using ultrasound tissue characterization

Manuscript Number SSMR-D-18-00067

Summary:
The study is to assess the inter and intra-rater reliability of UTC of patellar tendons not affected by BTB ACL graft and patellar tendons that have been used for BTB ACL grafts. UTC has the ability of characterising tendon tissues into 4 categories, Echo-type I (green) - generated by intact and aligned secondary collagen bundles, so called "fascicles". Represented by steady pixels with constant high grey levels on the echo pattern; Echo-type II (blue) - generated by discontinuous, swollen and wavy secondary collagen bundles, also called "fascicles". Represented by pixels that have some steadiness and variation of about 10% of the grey levels; Echo-type III (red) - generated by a loose matrix consisting mainly of much smaller fibrils. Represented by unsteady pixels with grey levels variation of more than 10%; Echo-type IV (black) - generated by mainly amorphous matrix with cells and fluid. Inter and intra-rater reliability was predominantly good to excellent except for volume which was poor to good.

The article has some good merit and with some minor editing and changes then it will be easier for the reader to follow.

COMMENT 1: The title doesn't translate to what the article was setting out to achieve i.e Inter and intra-rater reliability of UTC for harvested PTs

REPLY: We thought it was opportune to use the term Clinimetrics rather than intra and inter-rater reliability once Clinimetrics is described as “a methodological discipline with a focus on the quality of measurements in medical research and clinical practice. The quality of measurements includes both the quality of the measurement instruments and the quality of the performance of the actual measurements”.

We have no objection in being more specific in the title though.

ACTION: We have modified the title to: Reliability and normative values of quantitative assessment of harvested and unharvested patellar tendons of ACL injured athletes using ultrasound tissue characterization

COMMENT 2: Should the intext referencing be before the full-stop?

REPLY: We used the referencing system advocated by the journal (BioMed Central). If a different referencing system is also accepted by the journal, we would not have any issue in changing.

ACTION: none
COMMENT 3: What was the method of placing the skin markers, palpation or us? Explain.

REPLY: We used palpation to delineate the anatomical areas. The skin marks were used merely to facilitate placing the UTC tracking device aligned longitudinally with the patellar tendon.

ACTION: We added in page 4 the following: “Examiners used skin markers to delineate, by palpation, the apex of the patella, edges of the patellar tendon and tibial tuberosity. This step is not compulsory; however, it helps the examiner to place the tracking device longitudinally aligned with the patellar tendon regardless of anatomic variations.”

COMMENT 4: 600 sequential images would give a range length of 12cm, so I would guess only approximately 300 sequential images was actually used of the PT.

REPLY: Exactly, the amount of sequential transverse images included is related to the subject’s height. For instance, the longer tendon length we measured across the 37 subjects was 6.32 cm and the shorter tendon length was 3.38 cm, meaning within these participants’ tendons 316 and 169 transverse sequential images were considered respectively.

ACTION: none

COMMENT 5: The "Repeated Measures" paragraph is confusing. Be clear on what you mean by harvested, unharvested, unharvested uninvolved, ACL harvested.

REPLY: We agree with you. Thus, we have removed the terms involved/ unininvolved. Additionally, we changed the data presentation, bringing the number of tendons analyzed rather than number of subjects.

ACTION: We modified the text of “Repeated measures” (Page 5).

-we added a diagram describing the samples and statistical analysis performed (page 6).

COMMENT 6: Make the heading for the measurements in the methods and the results match, therefore it will be clearer what results match with how this was done in the methods. Easier for the reader to follow and they can easily refer back and forth from methods to results.

REPLY: Thank you.

ACTION: We modified the headings of methods, and results sections.
COMMENT 7: Explain what 'window size 17' is. Especially as you make reference to in the 'caution' chapter.

REPLY: The UTC imaging software has a few established settings that we can’t modify and others that we can chose among few options. The window size is one of these parameters. The examiner can choose among, 25, 17 and 9 window size.

Majority of the current literature has used window 25 in their methodology.

There is no explanation about the differences among them in the UTC imaging manual. However, after discussing this subject with the UTC imaging team, they explained that the lower the window size, more in depth you explore the tendon structure. The UTC software reads pixels and makes comparisons of sequential pictures based on how stead these pixels are. I understood as when using window 25 I have a zoomed picture, but when I diminish the window size I get the picture sharper; thus, I can see more in detail.

The window size selection is included in the caution chapter to avoid issues with comparisons. It is very likely that the proportion of type III and IV in our data is bigger than in other studies due to the difference in this parameter, thus comparisons between studies should take that into account.

Additionally, we have selected this setting on purpose, because we wanted to know in detail how the development of the healing tissue on the harvest site would evolve.

ACTION: We modified the text to:

“For instance, the tibial notch was not always centralized due to the harvested bone plug removed from the tibial tuberosity, and data of the distal portion of the tendon was included. Moreover, window size 17 was chosen for analysis for more detailed information of these harvested tendons rather than window size 25 as mostly adopted by the current literature. Thus, when comparing data of different studies, this difference in analysis setting should be considered.

Additional validation study could be performed since the original validated study was done in horses’ tendons and utilized different echo-types distribution than the ones utilized nowadays. Furthermore, despite the growing popularity of imaging modalities in clinical practice, the relative cost of UTC imaging might be a limitation.”

COMMENT 8: "different scans of same subject or same scans" I'm confused what you really mean, might be clearer with clearer headings as I've mentioned above, or use repeated measurements. Also be consistent with use of participants throughout the article.
REPLY: "different scans of same subject" means when we acquired data more than once of the same tendon.

“same scans” means that two examiners analyzed a single patella tendon data, twice.

ACTION: We revised the headings to facilitate understanding.

-We reviewed the number of participants in each statistical analysis, as there were repetitive subjects on the same sample due to acquisition at different times of their rehabilitation program. Thus, it is expected to have less participants than patellar tendons in each analysis.

COMMENT 9: "contour", is this like a region of interest box, as shown in figure 2? Or is it actual trace along the anterior surface of the PT, tibial tuberosity, posterior surface, apex and then back to the anterior surface?REPLY: Contour is the actual trace around the patellar tendon. However, a contour was drawn in each area of interest. Then additional contours were drawn along the tendon length to provide more information about the tendon to the UTC software.

- The order we draw the contours was a strategy to minimize errors. We started drawing a contour at the most distal point of the tendon (tibial notch), then another contour was drawn at the patella apex. Distance between these 2 contours equals the length of that tendon.

Areas of interest are fixed: (1) patella apex (2) proximal tendon, (3) mid tendon, and (4) distal tendon. We used the first two contours to determine the location of the other 3 areas of interest, as area of interest (2) equals 26 slices distally from area of interest (1), area of interest (3) equals 51 slices distally from area of interest (2), area of interest (4) equals 75% of the total tendon length.

ACTION: We added a Figure with examples of contours in harvested and unharvested patellar tendons (Fig 3.).

COMMENT 10: Line 126 - no need for commas either side of the brackets (reference mark 2)

REPLY: noted

ACTION: removed

COMMENT 11: Line 138 - comma after pattern

REPLY: noted
ACTION: included

COMMENT 12: Line 143 - no need for much
REPLY: noted
ACTION: removed

COMMENT 13: Line 148 - how is echo type IV represented?
REPLY: noted

ACTION: We included; “Echo-type IV (black) - generated by mainly amorphous matrix with loose fibrils, cells and fluid (hematoma and exudate). Represented by echoes with a severe lack of stability over sequential transverse images.”

COMMENT 14: Results: Same note again about matching headings with methods so that it's easier to follow.
REPLY: Noted.

ACTION: We modified the headings in all sessions to facilitate understanding.

COMMENT 15: Why are the length measurements not included in the tables?
REPLY: The tables have already a lot of columns, so one reason was space.

Another reason is that the length is not related to any area of interest in specific because includes all of the areas, so we could not visualize a nice way to include this data in the existing tables.

We thought it was unnecessary to create a new table only for the tendon length.

ACTION: none

COMMENT 16: In the tables, what are the units of measure for the echo types, volumes, thicknesses?
REPLY: noted
ACTION: We included % for the echo-types, cm3 for volume, and cm for thickness. Also, we explained that each pixel unit can be considered as equivalent to 1cm (page 5).

COMMENT 17: Table 1 and 2 have the same headings even though showing different data. Same comment as with consistency of matching headings as above.

REPLY: noted

ACTION: We changed the titles of tables and presented the “n” as number of patellar tendons included in the statistical analysis rather than number of subjects.

COMMENT 18: Table 1 and 2 - how did you get n=10 and n=20 from 19 participants?

REPLY: From the comments we received we decided to review the data presentation. So instead of using the “n” as number of subjects, we are now using “n” as number of patellar tendons included on each statistical analysis.

ACTION: We changed the titles of tables and reviewed the number of subjects and patellar tendons included on each different statistical analysis.

COMMENT 19: How is the thickness and volume measured? Should be in methods.

REPLY: noted

ACTION: We included: Measurement of thickness of the mid tendon was done manually by the examiners by activating the measuring tool of the UTC imaging software. (Page 7)

“In addition to that, the UTC algorithm also quantifies the area within the contour (volume) drawn on the selected areas of interested (reference marks 1-4).” (page 8)

COMMENT 20: Line 191 - why analysing the data as mixed harvested and unharvested? Why not separated?

REPLY: Because we wanted to know how reliable it was the acquisition and the analysis of patellar tendons between more experienced and less experienced examiner regardless if the tendon was harvested or not. Thus, the same patellar tendon was acquired and analyzed by 2 examiners. We used 11 harvested and 10 unharvested tendons for this statistical analysis.

ACTION: none
COMMENT 21: Table 3 and 4 - same points as above, re headings/units of measure. Again n=20 from 27 participants?

REPLY: Same explanation as above.

ACTION: We reviewed and modified the data presentation.

COMMENT 22: Line 211 - surprising inter-rater reliability was poor for volume when examiners processed the same scans, but better reliability when examiner did their own scans and then analysis, I would have guess the extra step would have resulted in poorer reliability.

REPLY: When any examiner draws the contours, it is advised that you draw them on the very edge of the tendon, ensuring to be within the limits of the tendon.

This method ensures that you select and feed the software with exclusive tendon tissue. The variable of volume is the one to be less reliable between different examiners because you one examiner can be farther or closer to the edge of the tendon. The proportion of aligned tissue within smaller or greater area will remain the same, however the area within the contour that represents the volume can differ easily.

ACTION: We included explanation of the method to calculate volume and modified the discussion (page 14)

COMMENT 23: Discussion: The start of the discussion is just a summary of the results, and the first paragraph is then just repeated below in the next couple of paragraphs.

REPLY: The first paragraph of the discussion usually summarizes the main findings of the study which are discussed later on.

ACTION: Slightly modified, no further action taken.

COMMENT 24: Line 253 - Due to consistency? Is this right?

REPLY: noted

ACTION: We modified this part of the discussion. (page 14 & 15).
COMMENT 25: 258 - ?excellent, not for all measures.

REPLY: Thank you, you are right.

ACTION: We modified the text as presented in page 14

COMMENT 26: 265 - "characterise the quality" this wasn't assessed.

Reading through the references, it appears most of the articles have used:


as the gold standard for the 4 tissue characterisation types, although this study was on horse tendons and had a sample size of 4 (2 normal and 2 abnormal). Do you know if this has been further validated on human tendons?

REPLY: Validating studies were done in horses and to the best of my knowledge are the studies justifying the use of this method in humans.

There is one study testing the accuracy of the UTC imaging method to discriminate symptomatic vs asymptomatic Achilles tendons (Br J Sports Med. 2010; 44:1153–9). They found an accuracy of 83% in allocating the tendons in the right group.

ACTION: We added in the caution chapter that further validation study in human tendons would be appropriate.

COMMENT 27: * Figure 1 - what's the difference between C and D?

I'm not an expert in UTC but am with clinical US. Looking at the images shown in the article the gain or amplitude/power used appears to be set too high. This may be normal for UTC and I don't know the answer to this, but wouldn't having the data collected "over gained" affect the characteristics as too many pixels will be at the maximum grey scale intensity of 256? How is the gain adjusted or levelled correctly? Or is it just the quality of the image in the PDF?
REPLY: There is no difference between figure 1C and 1D. They just appear concomitantly in the UTC acquisition and analysis window. We wanted to show the readers exactly the window view of the UTC software.

Regarding the settings there is not much an examiner can do to change that since they are pre-established. One can change the depth and focus by changing the” size” of the subject.

Some studies have used small size subjects at their settings, which results in reduction of depth to 3cm and focus at 1.3cm. we have tried to use the same settings, but for us, it wouldn’t work because the knees with harvested tendons can present big swelling, and the focus would be placed at the subcutaneous tissue rather than tendon.

ACTION:none

Marijke Welvaert, PhD (Reviewer 3): Upon request from the Editor, I focused my review on the statistical aspects of the manuscript. Overall, I think the statistics are well reported, the chosen method is appropriate and the interpretation of the results support the presented numerical outcomes.

COMMENT 1: Main comment: One aspect that is missing from the manuscript is a discussion on the assumptions underpinning an ICC analysis. The most straightforward one is the normality assumption and while there is a certain robustness against violations of this assumption, the reader should have information around the distribution of your data.

REPLY: We thank you for your comment.

ACTION: We have examined all the data for normality through graphical estimation and calculation of Kolmogorov-Smirnov and Q-Q plots. The majority of the data were normally distributed for all the parameters estimated and given the opinion of Hays* we suggest that for these analyses ICC ANOVA was appropriate.

COMMENT 2: A couple of very minor comments:

- l. 159 Please add the confidence level (e.g. 95%)

REPLY: Noted.

ACTION: We added in page 8. “Confidence interval (CI) of 95% was adopted. Its values will be presented between parenthesis after the ICC values”.

COMMENT 3: - Tables: Please add units of measurements either in-table or as part of the explanatory notes to be consistent with in-text reporting.

REPLY: Noted.

ACTION: We added in all the tables headings the units of measurements.