Reviewer's report

Title: Repeatability and agreement of ultrasonography with computed tomography for evaluating forefoot structure in the coronal plane

Version: 0 **Date:** 15 Dec 2016

Reviewer: Lisa Newcombe

Reviewer's report:

Reviewers Report

Manuscript Title: Ultrasound evaluation of forefoot structure in the transverse plane: a validation study

Thank you to the authors and the editorial team for the opportunity to review this interesting article, which provides a novel and valuable contribution to the field. Overall, the article is reasonably well written and many of the decisions surrounding the methods are justified. There are some clear areas where improvements can be made with some minor amendments. However, I feel that there are some significant issues that need to be either clarified in detail, and/or addressed with major amendments prior to being suitable for acceptance for publication. I have provided my point-by-point responses below and have provided an annotated PDF to assist with the review and revision processes.

1. Is the question posed by the authors new and well defined?

Yes, but perhaps there could also be more clarity in the definition of 'forefoot structure' to differentiate this work, investigating forefoot plantar soft tissue thickness, forefoot transverse arch height/length/index, sesamoid angle, from ultrasound work investigating inflammatory features in metatarsophalangeal and interphalangeal joints in rheumatology.

2. Are the methods appropriate and well described, and are sufficient details provided to replicate the work?

The methods are appropriate, but would benefit from further refinement and reference to the appropriate literature concerning the key outcome measurements assessed. Sufficient details are not provided to replicate the work in all areas of the methods, please see my comments relating to the methods under 'Major Compulsory Revisions'.

3. Are the data sound and well controlled?

The data appears sound and well controlled, however, confounding variables could be better considered in this analysis. The band and Altman plots to the best of my knowledge appear to be analysed and presented correctly. However, I am not an expert statistician and draw on my knowledge, skills and experience as a clinical researcher.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?

Yes, but there is a statement that some data is not shown, relating to information collected on the presence of each participant's forefoot structural deformities (i.e. hallux valgus, lesser toe deformities) which I feel the article would benefit from including, as deformities are a major confounding variable to results (acknowledged in the background section), which is not acknowledged in the results or discussion, or when attempts are made in the discussion to compare outcome measure values with values found in other studies.

5. Are the discussion and conclusions well balanced and adequately supported by the data?
The discussion is reasonably well-written but there are some statements and conclusions which are not supported by the data or literature or are discussed without sufficient consideration of the impact of bias introduced by the study limitations. Additionally,

6. Do the title and abstract accurately convey what has been found?

To an extent yes, but there are some inconsistencies and the conclusion of the abstract needs re-wording which is consistent with comments relating to the article conclusion, which requires more acknowledgement of study limitations with relation to overall conclusions. From the title and abstract, it isn't immediately apparent what type of validity is being assessed, which appears to be concurrent validity, where CT is the accepted reference standard. As there are other forms of validity that are not measured, this should be made clearer in the title and abstract. More information could also be included in the abstract methods including details of the researcher, ultrasound and CT equipment, time points studied, participants being healthy volunteers. It is not clear from the background in the abstract why US and CT are being compared specifically and the parameters of forefoot structure being referred to. In the results section, almost perfect agreement is reported for all parameters for ICC and Pearson correlations, which should be amended in the abstract as not supported in the article results section.

7. Is the writing acceptable?

Yes, overall the writing standard is good. There are some minor spelling mistakes and grammatical errors which I have highlighted, and some suggestions of re-wording sentences for clarity.

Minor Essential Revisions

* Background, p3, paragraph 1, line 80: Reduced plantar tissue may be associated with lesser toe deformities across a range of age groups, not restricted to the older population, however, I appreciate the referenced work has demonstrated this only in the older population. However, this may be a point to consider in the population studied in this research.

* Background, p3, paragraph 3, line 92: When ultrasound is introduced as a diagnostic imaging modality, perhaps it would also be useful to add in 'musculoskeletal ultrasound'

* Background, p3, paragraph 4, lines 100-102: More clarity could be provided on the established techniques for measuring the outcome measure parameters stated and their limitations, as well as details on why these variables are key to measuring forefoot structure statically, with links to foot function, where appropriate, i.e. detailing the significance of the transverse arch index. This is provided in the discussion, but could be introduced earlier for better understanding.

* Background, p3, paragraph 4, line 97:

* Please consider reviewing and adding in the following key references into this paper for this section.

Cavanagh, P.R. Plantar soft tissue thickness during ground contact in walking. Journal of Biomechanics. 1999; 32:623-628.

* It may also be useful to detail the use of ultrasound to measure plantar soft tissue thickness in the foot has been cited by other studies including the following study:

Telfer S, Woodburn J, Turner DE. Measurement of functional heel pad behaviour in-shoe during gait using orthotic embedded ultrasonography. Gait & posture 2013: 39(1):328-32

* Methods, p4, paragraph 2, line 119: A brief description of what ImageJ software is would be useful.

* Results: p5, paragraph 2, line 150: Could add in the ICC result for the AUTA parameter in text.

* Results, Table 1, line 158. Are these measurements mean (SD) measurements? This should be stated.

* Discussion, p7, paragraph 3, lines 189-194: Could this information be described earlier in the background/methods sections to provide clarity earlier of the selection of these key variables, as really useful to understand the need for these key variables to be included in the data collection?

* Discussion, p7, paragraph 3, lines 196-197: In consideration of this statement, it may be useful to note it is not yet known how these parameters compare with clinical assessment methods, which based on these study findings, could be tested in future work, also assessing dynamic changes during gait.

* Discussion, p7, paragraph 3, lines 203-204: Could more information be described as to why female health volunteers was a limitation in terms of the population studied, and the reason for this recruitment pragmatically (taking into account sampling technique and generalisability to the overall population)? This is important where forefoot and toe deformities can also exist in this population.

Major Compulsory Revisions

* As an overall comment, the type of validity being assessed isn't immediately apparent as concurrent validity from the title, abstract or conclusions. As there are other forms of validity, this could be expanded upon and made more explicit in the title and conclusions.

* Background, p3, paragraph 1, lines 78 and 79: The referenced study (9) does not support this statement as this study found an inverse correlation between plantar thickness and increased plantar pressures, indicating tissue thickness decreased with high plantar pressures. This also varied with the presence of callous. Please review this paper and revise this sentence.

* Methods, p3, paragraph 1: For this study to be reproducible, more detail is needed to explain the inclusion and exclusion criteria applied. It is not clear why only women were recruited, and

which sampling technique was used. It is not clear if any of the participants in the study suffer from any of the foot disorders explained in the background information, i.e. diabetes, lesser toe deformities. More demographic information could be described (such as BMI which could impact on findings as is not detailed as controlled, or as a limitation) and more clinical information such as the presence or absence of forefoot structural changes (i.e. hallux valgus, lesser toe deformities introduced in the background information) could be described as confounding factors. It might also be useful to provide details on the hospital/institution recruited from.

* Methods, p4, paragraph 2, line 112: It would be useful to detail the professional role of the person undertaking the ultrasound scans and their level of training in diagnostic ultrasound (clear that this training was for 3 months, but the level not clear, which could be expanded upon as a potential limiting factor in the discussion).

* Methods, p4, paragraph 2, line 112: Good detail has been added about the type of transducer utilised, but unsure if the exact model of the ultrasound machine has been cited, only the manufacturer name as Hitachi Aloka Medical. Whilst the frequency probe range was specified as 5-10 MHz, details of standardisation of other B-Mode settings are important to be cited for the purpose of validity and repeatability testing which are currently not cited. I.e. standardisation of gain, depth, frequency, focal zones maintained for both ultrasound scans, and between participants.

* Methods, p4, paragraph 2, lines 115: More detail could be added in for the exact anatomical landmark the lines were drawn from on the 1st and 5th metatarsal heads. For example, was this in the centre, from most medial to lateral plantar aspects, distal or proximal aspects? Was this standardised?

* Methods, p4, paragraph 2, lines 117 and 118: More information could be added in description of the placement of the transducer in the appropriate scanning plane. 'Transverse plane placement/orientation of the probe' should be added, alongside details of scanning on the line connecting the metatarsal heads. This detail is present in figure 2, however.

* Methods, p4, paragraph 2, lines 117 and 118: In consideration alongside Figure 2a, were the metatarsophalangeal joints held in a neutral, dorsiflexed position the researchers non scanning hand?

* Methods, p4, paragraph 3, lines 121-125: It could be made clearer how many CT images were obtained, as was detailed for US. It could also be made clearer why US and CT are being compared directly, and not X-ray or MRI. For both CT and US, detail should be added that scans obtained in an unloaded position.

* Methods, p4, paragraph 4, lines 127-135: More information relating to the description of the exact placement of the measurements during the image appraisal could be detailed. This is to provide clarity on which anatomical location of the metatarsal heads marked 'the bottom'. Was this the lowest point of the epiphysis? I am also unclear of what is meant by the 'six plantar points at the shortest distance from the six bone points', more clarity to be provided for reproducibility. Additionally, it could be made clearer that one researcher placed and measured these points, without blinding (which should be recognised as a limitation in the discussion, unless blinding was undertaken but not noted, see below comment).

* Methods, p4, paragraph 4, lines 127-135: More information could be provided on the significance of the 2nd MTH length measurement (i.e. why weren't all MTHs lengths measured?) although from the back ground, this could be understood as an area of increased shock absorption.

* Methods, p4, paragraph 4, lines 127-135: Two researchers trained in ultrasound could have reviewed both CT and US scans blinded for better measures of inter-observer reliability, and measures of repeatability. Although, recognised as a limitation in the results.

* Methods, p4, paragraph 4, lines 127-135: Perhaps undertaking 3 ultrasound scans instead of 2 scans would have improved the meaningfulness of an average for the ultrasound data.

* Methods, p4, paragraph 4, lines 127-135: Researcher blinding between the CT and US image appraisals has not been discussed and inter-observer repeatability, ideally between time points, utilising standard ultrasound settings would have improved the reliability and repeatability measures of the study, which should be clear in the discussion section (as is mentioned for inter-observer assessments).

* Discussion, p7, paragraph 1, lines 163-164: This statement also needs to acknowledge this result is based on intra-rater agreement, of 2 US scans, taken at a single time point.

* Discussion, p7, paragraph 2, lines 170: Please describe the other methods referred to, as not detailed in background information.

* Discussion, p7, paragraph 2, lines 171-176: Please also detail other confounding factors on differences between results, which may include/exclude standardisation of ultrasound pre-settings and operator dependent B-mode settings (if not standardised). Potential influence of operator scanning technique if metatarsophalangeal joints held neutral/dorsiflexed, compared to CT scans. This may have had a larger impact on patients with lesser toe deformities, but no information has been provided on forefoot structure of participants.

* Discussion, p7, paragraph 2, lines 181-182: More detail could be added of other potentially confounding variables including the demographics of the participants, i.e. BMI, forefoot structural deformities, medical conditions/history, i.e. inflammatory arthritis, diabetes.

* Discussion, p7, paragraph2, lines 188: Information on the presence of hallux valgus in 16 participants is introduced in the discussion with data not shown. Details on this data could add understanding of the confounding variables impacting the other parameters detailed in the results.

* Discussion, p7, paragraph 3, lines 200-201: 'Since we targeted the forefoot bony structure, the inter-observer reliability of our US methods may as good as the previous study'. I am not sure this can be supported by the information provided in the study as there is missing detail on exact placement of measurement points.

* Discussion in general: Some good limitations acknowledged but there are several sources of bias in this study which need to be addressed such as:

o No blinding between appraisal of US and CT images discussed

o Impact of US settings unclear. A statement is required in methods to state the standardisation of B-mode US settings, whether this was present or not, and if not, the impact of this on operator dependency and results of validity and repeatability. This may also apply to CT scans between participants.

o Data collected at a single time point, and repeatability not measured for the same participant at two separate time points, with blinding between days.

o Other confounding variables must also be acknowledged relating to the participants demographic data and data omitted related to forefoot structure.

o Good acknowledgement of intra-observer repeatability and validity measured only between CT and US measurements, but perhaps this could be emphasised more in relation to the recommendations made in conclusions and the discussion

Please consider and discuss all potential sources of bias and confounding.

* Other limitations to be discussed include the sample size, in line with the statistical analyses undertaken, with details of a statistical power test to determine sample size if appropriate.

* Conclusions: In terms of the information provided in the methods and results to support the conclusions made, I think it needs to be clearer that (with acknowledgement of the study limitations), this study provides new evidence of concurrent validity for the outcomes measured (giving a better insight into forefoot structure in general, but not all measurements of forefoot structure) between US and CT, and not validity in general. There is therefore potential for US to be utilised in future work investigating forefoot structural changes, but more work needs to be undertaken in future to investigate inter-observer reliability, to address bias limitations and to make comparisons with clinical assessment, before this method is recommended for foot screening and as an alternative for other forefoot structure measurements (statically, as not assessed dynamically). I recommend rewording the conclusion and the abstract conclusion to reflect this.

Discretionary Revisions

* Background, p3, paragraph 3, line 94: This sentence reads a little unclear and could be re-written to state the non-invasive and non-ionising advantages of US compared with other imaging modalities.

* Methods, p3, paragraph 3, line 121: Participants lay in 'a' supine position.

* Line 325: Spelling mistake as follows: 'metatarsopharangeal' joints

* Line 337: of two 'of' the red lines

* Lines 177-178: showed 'the' 2ndMTH

* Discussion, p6, paragraph 3, lines 194-196 Please revise the following sentence. 'Because soft tissue thickness is related to diabetic foot and toe deformities, AUTA would provide better insight into these foot problems'. This sentence could be re-written to state that better accuracy in the measurement of the AUTA may offer better understanding of changes which can occur in patients with structural forefoot deformities.

* Discussion, p7, paragraph 3, lines 204: employees

Level of interest
Please indicate how interesting you found the manuscript:

An article of importance in its field

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal