Author’s response to reviews

Title: The ankle brachial index in people with and without diabetes: Intra-tester reliability

Authors:
Sarah Casey (sarah.casey@newcastle.edu.au)
Sean Lanting (sean.lanting@hotmail.com)
Vivienne Chuter (vivienne.chuter@newcastle.edu.au)

Version: 1 Date: 30 Mar 2020

Author’s response to reviews:

Reviewer reports:
Reviewer #1: This is a well-conceived and important piece of research, which provides valuable clinical information on the reliability of Ankle Brachial Indices (ABIs) in a specific population (i.e., people with diabetes). The researchers recruited a relatively large sample and have taken care to design a high quality experiment. I offer the following comments for the authors consideration as they finalise the paper for publication:

Major Compulsory Revisions
1. Consider removing the discussion on predictors of low ABI and reporting these results in a separate paper if there is sufficient data to do so. This is a sizeable topic on its own, which is relatively disparate from measurement reliability. Given both areas require introduction of relevant literature and a dedicated discussion on methodology, results and findings, it would seem more fitting for the respective topics to be addressed in stand-alone publications. The lack of participants presenting with low ABIs also raises questions about statistical power, in addition to other methodological considerations about that part of the study. For example, I query whether an ABI of .99 can be considered borderline, particularly when accounting for measurement error. Separating the topics into different papers would allow greater space to explore each area in more detail, including further explanation and elaboration on the reliability study findings in this current paper (see suggestions below).
   Thank you for your comments. We have removed the analysis of factors contributing to ABI variability.

2. The authors have done very well in presenting the data however the narrative describing the reliability of ABIs reads as inconsistent at times. Specifically, the ICCs are described as excellent, however interpretation of the absolute measures such as the LOA’s, MDC and the 95% confidence intervals of the ICCs, suggest a more moderate picture (for e.g., the lower end of the 95% confidence interval for ABI reliability in people with diabetes is 0.62). Consider altering the descriptors used for the reliability coefficients, clearly acknowledging that the descriptors are an indicative term only, with the interpretation of the ICC depending on several factors, including the proportion of the overall measure which can be attributed to error and how the results are to
be applied clinically. For example, instead of 'excellent reliability' for an ICC &gt; .75 as used by Fleiss (1999), Portney and Watkins (2009) use the term 'good reliability' for the same coefficient. In addition, when discussing ICCs, remember to reflect the 95% confidence interval as an indicator of precision of the ICC. Ensure this is then contextualised against the indicators of absolute reliability, for a well-rounded discussion. Whichever way you decide to address this issue, ensure there is a cohesive, consistent and accurate approach to interpreting the reliability data, so that readers understand the position you are taking on the measure (also see points 3 & 4).

Thank you for your insight. We have followed your suggestion and amended our interpretation of the ICC values in line with those suggested by Portney and Watkins for use in clinical research. The ‘Statistical Analysis’ section has been amended to reflect this change and now reads: ‘All ICC values were interpreted according to cut-offs suggested by Portney and Watkins, in which &gt;0.75 denotes good to excellent reliability; 0.50 to 0.75 suggests moderate to good reliability, 0.25 to 0.50 denotes fair reliability and 0.00 to 0.25 denotes little or no relationship’.

The Discussion has also been amended:’ The intraclass correlation coefficient (ICC) values obtained results suggest that the intra-rater reliability of the ABI is good, with values that are comparable to the findings of previous studies in mixed populations including those with risk factors for and/or suspected PAD…”

3. The authors have done an excellent job in operationalising the reliability co-efficients (i.e., ICCs), through the use of statistics reporting absolute measurement error in the units of measurement, i.e., LOAs, SEMs and MCD. It would be useful for the readers to have each of these statistics clearly explained in simple terms however, including why the three measures were reported, how the measures differ and what each adds clinically. Specific examples of how each may be applied clinically, by applying the data to a scenario, would be very useful. I suggest selection of a mixture of best and worst case scenarios for these clinical examples, so that readers gain an appreciation of the range of error that needs to be considered and how the error can be quantified in practical terms based on a given ABI measure.

Thank you. The Discussion has been amended to try to better explain the interpretation of LOA ‘…it is noteworthy that the LOA were quite large with upper and lower bounds ranging from -0.16 to 0.16. This indicates that the ABI may exhibit measurement variation that affects its ability to distinguish true change from measurement error.’ SEM and MDC are described as follows: ‘Similarly, the SEM was 0.03, which suggests that the variability of the test means that a large change in ABI is required clinically for there to be certainty around it being a true change rather than variability of the measurement. The MDC, which reflects the amount of change required to indicate an effect above the variability of the test, was found to be 0.08 for the ABI overall, and in participants without diabetes.’ An example for MDC is provided, however and additional sentence has been provided, so the section now reads as follows: In practice, a decrease of 0.11 in ABI between measurements may represent a change from a normal value of 1.00 to 0.89, indicating a pathological value, however this may occur as a result of measurement variability rather than disease. Similarly, a value of 1.4 may incorrectly indicate the presence of MAC in a person with a true ABI of 1.29.’

4. Be specific when describing the clinical implications of the work. On one hand the paper states that reliability is excellent however on the other hand the results should be interpreted with
Minor Essential Revisions
5. Consider describing the degree to which the clinician taking the measures has past experience taking ABIs on the population studied,

Thank you. We have amended the sentence to include more detail: ‘a Podiatrist with more than 15 years’ clinical experience, using ABI several times a week.

6. Ensure that minimal detectable change is not conceptually mixed up with minimal clinically important difference (see line 79 in the background),

Thanks for pointing that out. The sentence has been amended: ‘There are currently little available data quantifying the magnitude of error that occurs with the measurement and therefore the amount of change in ABI that represents a true change (over and above measurement error).’

7. Please clarify the inclusion criteria regarding (line 98). It is unclear whether you recruited participants with pre-existing PAD using ABIs or some other approach. If PAD was an inclusion criteria, make it clear early in the paper that people with PAD, and people with PAD with diabetes, were recruited. Or are you saying that people who met the guideline for testing for PAD with ABIs were recruited? Either way, please double check the wording as I read this a few times to try and establish the inclusion criteria that were used. Also clarify, is it people 65 years and over were included, OR or 50 years and over if they had a history of diabetes?

The sentence has been amended: ‘… 65 years of age and over, or aged 50 years and over with risk factors for atherosclerosis (e.g. diabetes mellitus, history of smoking), or in individuals with exertional leg pain or non-healing wounds.’

8. Reference the standardised ABI techniques followed in the methods

Thank you. An additional sentence has been added as follows: ‘The test procedure was carried out in accordance with the American Heart Association Scientific Statement’, with the relevant reference.

9. Line 61 - add a reference for the statement "The ankle brachial index (ABI) is typically calculated as the ratio of the highest of the dorsalis pedis and posterior tibial artery systolic pressure to the highest of the left and right brachial systolic pressures".

Thank you for identifying this oversight. A reference to the AHA Scientific Statement has been added.

10. Consider adding to limitations, one person only assessed ABIs only therefore generalisability to other clinicians may be limited, that the clinician was not blinded to the measures therefore may remember some patients measures (especially unusual measures) and that using ABI to
recruit participants may involve some inaccuracy (due to the measurement error you went on to establish).

Thank you for your comments. The section has been amended to read: ‘. A further limitation is the lack of blinding of the clinician to the pressure values obtained. This study only established the intra-rater reliability of the ABI when assessed by an experienced podiatrist. In clinical practice, the ABI is used by a range of health professionals, and as such the findings of this study may not be generalisable to other clinicians. In addition, the inter-rater reliability of the ABI is of considerable importance, considering that it is often used by different personnel rather than always being repeated by the same clinician.’

To clarify, participants were not recruited on the basis of diagnosis of PAD vs no PAD using ABI. Hopefully the manuscript adjustments made in response to your earlier comments have clarified the methodology.

11. You have some very rich data therefore consider exploring it in more depth. Some ideas, to take or leave as you wish, are: that the brachial systolic measures are less reliable than the ankle systolic measures - what might this mean to the ABI reliability results and might it be an area to address the improve reliability of the coefficient?

Thank you for your comment. We agree that the lower reliability of the brachial pressures are likely to reduce the reliability of ABI and should be targeted to try to improve overall test reliability, and we have addressed this as per the below:

‘Overall brachial pressure measurements were found to be less reliable than ankle pressure measurements in participants with and without diabetes, which is consistent with previous research using handheld Doppler in podiatry practice. The use of automated oscillometric systolic blood pressure measurement has been shown to be more reliable than manual Doppler method in people with diabetes. Given that ankle pressures were found to be highly reliable (ICC 0.88-0.92) improving the reliability of brachial pressure measurement is likely to improve reliability of the ABI overall. Furthermore, the use of automated pressure measurement may reduce the time taken to carry out ABI, which is one of the most frequently cited reasons not to perform vascular assessment.

How might the differences between groups in age and BMI be relevant?

We have updated the discussion to address this: ‘Diabetes has been associated with several complications that affect diagnostic accuracy of the ABI for PAD including medial arterial calcification (MAC). It is possible such complications could affect reliability, however to our knowledge this has not yet been established. The diabetes cohort recruited to our study, although having increased risk of PAD (e.g. higher BMI, diabetes, insulin use), did not demonstrate overt clinical signs or symptoms of PAD and MAC. This cohort also had a relatively short mean duration of diabetes (&lt;10 years), making them less likely to have developed related complications. Similarly, the no diabetes group were significantly older (mean age &gt;75 years) placing them at greater risk of age related PAD and MAC but lack of abnormally low or high ABI results suggest this was not the case. This may have resulted in the similar reliability of the test in both groups.’

Did some participants have wounds and was this even between groups? How might this be relevant to the study findings?

No participants had wounds
Why was there only a small difference in ABI reliability between the diabetes and non-diabetes participants? Was this what you expected and how do you explain this finding?

Thank you. This has been added to the discussion in response to your earlier comment relating to BMI and age.

What are the implications of the study in term of developing the ABI measure - is there room to improve accuracy, what are the likely sources of error and can some be worked on?

Possible sources of error are described in the fourth paragraph of the Background, lines 71-75. As a reliability study, we have also discussed the use of automated brachial pressure cuffs to improve reliability of this measure and overall reliability of the ABI (lines 262 – 268).

And finally, I read it that you undertook paired t tests on time 1 measures v's time 2 measures. What was the purpose of these comparisons and what is your interpretation of the result? Where you looking to establish systematic versus random error? What did you find?

The paired t-tests were performed to check for systematic differences in the test and retest results as per the discussion: ‘There was a statistically significant difference in ankle pressure and brachial pressure values obtained between the first and second visit, however this was not reflected in a significant difference in the ABI result obtained. This finding may result from variation in the exact timing of the second appointment compared to the first (e.g. early in the morning for visit 1, mid-morning for visit 2), or it may be due to familiarity with the procedure, location and researchers at the second visit.’

12. Consider acknowledging the broad clinician concern that exists, questioning the validity of ABI's in diabetes (due to medial calcific sclerosis) as an important issue, albeit not what is being investigated in this study. One sentence would suffice, but it just shows an appreciation of the scope of accuracy issues in the area, for completeness.

Thank you for your comments. The Background section describes the potential sources of error when using the ABI in people with diabetes: ‘In diabetes cohorts, PAD is known to have a more aggressive disease presentation, to be more likely to affect infra-genicular arteries and to co-exist with medial arterial calcification which can prevent compression and pressure measurement of the lower leg arteries. Although there is significant evidence that these factors affect diagnostic accuracy of the ABI in this population’


Thank you for your thoughts. We refer to our systematic review, which included this paper as well as others that report inter- and intra-rater reliability.

14. Provide reference for line 79 statement "There has also been limited investigation of the intra tester reliability of the ABI in people with diabetes, with only one previous study, which reported a coefficient of variation of 8%.

Apologies for that oversight – the citation has now been included.

Discretionary Revisions

15. In line 154, would the wording "no ABI's...." read clearer than "not ABI's...."?
As part of the compulsory revisions requested, this line has been deleted.

16. Note, there are 2 full stops in line 195
Thank you for identifying this typo – this sentence has been removed during the compulsory revisions requested.

Reviewer #2: Many thanks for submitting your manuscript outlining your research.
Whilst I commend your work and efforts, I feel that in its present format, it is not an easy paper to read and for that reason, you may find many of your [intended] audience will not reach the end of your paper.
To assist you further, I would like to make the following comments and recommendations to support you in achieving publication:

Background:

In your second sentence of the opening paragraph, line 54, insert ’the’ between ”arteries,” and ”presence”. At the end of this sentence please clarify which population you are specifically referring to.
Thank you. We have made the suggested insertion and amended the last phrase to read: “…in people with diabetes”.

Lines 63-64 - you refer to a normal value of 1.00-1.40; 0.9-0.99 as borderline with below 0.91 representing PAD. Your reference here is surprisingly from a cardiology journal - I think there are more appropriate references you could use with more widely accepted reference values. The American Diabetes Association cites normal reference values as 0.91-1.3. Perhaps some discussion as to why these are different. What is the most current consensus document? This study investigates ABI reliability in the general population, with a subgroup analysis of people with diabetes. For this reason, we have referred to guidelines that include the whole population. The ACC/AHA guidelines reflect the importance placed on the ABI for its ability to predict the presence of atherosclerosis in cardiac and cerebral arteries. At the present time, there is not an international consensus document for PAD screening in the general population that covers all continents. The values used in this paper reflect the nature of the ACC/AHA committee who have updated the PAD guidelines in response to emerging research. Both ACC/AHA and TASCII suggest an upper limit of 1.40 for normal, with values above that representing the presence of MAC. The sentence has been amended to read: ‘A normal value for ABI assessment within the general population is considered to be 1.00 – 1.40…’, with a citation of the TASC II guidelines (lines 64 – 65).

Line 66 - you cite ”Current international guidelines....” but only give 2 references, one again, from a cardiology journal and another paper from 2007, which may not be thought of as current. Have you considered the IWGDF 2019 guidelines on PAD (https://iwgdfguidelines.org/pad-guideline/) and also A Systematic Review for the screening for peripheral arterial disease in asymptomatic patients, J Vasc Surg. 2015 Mar;61(3 Suppl):42S-53S for further authority? The conclusion from the 2015 systematic review does not support the benefit of routine ABI screening, and it would be good to hear some further discussion regarding the value of it in clinical practice. I do note however, that you refer to your own 2019 systematic review (reference 19) but I think readers would value some more background here to understand why
your primary aim was to just determine intra-tester reliability (and not to consider inter-tester reliability which would perhaps be more applicable to every day clinical practice, and has been reported as having little variability?). It might also be helpful for the reader to remind them of the difference between intra-tester and inter-tester variability.

Thank you for your perspective. We have cited published guidelines that are intended to influence clinical practice regarding assessment for peripheral arterial disease, for all sections of the population, including, but not limited to, those with diabetes. The TASC II Guidance has been reviewed, but not altered since 2007, so is still current. We agree that there is room to critique current guidance (with at least one published paper to our knowledge doing so), however we feel that is outside of the scope of this paper, which has as its focus the intra-tester reliability of the ABI.

This study was devised as a result of the findings of our systematic review, which found a dearth of high-quality studies investigating intra- and inter-tester reliability of the ABI. For this study, we chose to evaluate intra-tester reliability, and the study was set up to do that. The Methodology section clearly describes the process used for collecting data, which should enable the reader to identify that the study uses one tester who took both test and re-test measurements.

Methods:

Lines 105-107 Does this achieve adequate power? Many readers, not actively involved in research, will not be familiar with how you have presented this.

Thank you, the sentence has been amended as follows: 'In order to ensure adequate statistical power, a sample size of a minimum of 40 participants for each group (diabetes and no diabetes) was determined…’

Statistical analysis 130-162:

You have given nearly 10% of your paper to this section and I feel that you may lose all but the most hardened researchers and statisticians here. Whilst I do not wish to undermine the academic nature of your study (fitting for a research journal such as this), personally I think this section would be better summarised and simplified.

We understand your perspective. As part of the compulsory revisions required by Reviewer 1, the regression analysis has been removed, and further explanation of the statistical tests used has been included. Hopefully this will make the statistical analysis section more comprehensible to a broad readership.

Discussion:

Lines 202-203 you state "The results suggest that the intra-rater reliability of the ABI is excellent, which is comparable to the findings of previous studies......" but on page 4 under Background line 79-80 you state "There has also been limited investigation of the intra-tester reliability of the ABI in people with diabetes, with only one previous study....". Please could you address this inconsistency?

The sentence in line 202-203 describes our findings for the cohort as a whole, with references for those papers that investigated ABI reliability in a mixed population. The sentence you refer to in the Background section reports the existence of only one previous study that investigated ABI reliability specifically in people with diabetes.
I feel there are too many abbreviations in your discussion section - I would suggest either avoiding the less familiar abbreviations or consider including a list of abbreviations that the reader can easily and quickly refer to. Thanks for your feedback. The Discussion has been amended and now states the full term and its abbreviation when first used in this section, for all statistical terms.

Line 250 you note that the presence of diabetes was not significantly associated with an abnormal ABI. It is well documented that diabetes can be associated with falsely elevated readings and I would like to see some discussion of this here. This section has been removed as part of the compulsory revisions required by Reviewer 1.

You have not specifically included a section on 'Limitations' of the study. Lines 254-262 perhaps touch on this but I think it would be helpful to include a specific sub-section to discuss the study limitations. Thank you for this, we have included a limitations paragraph which we think is consistent with the journal format however happy to include this as a subheading if required/allowed.

Conclusions:
I don't like the use of abbreviations LOA and SEMs (line 266) - I feel these should be written in full. Thank you. The terms have now been stated in full.

It is not obvious what the implications are, of your research, on actual clinical practice. Could you include a summary here? An additional sentence has been added as follows, to reflect the summary provided in the Discussion: ‘As such, our findings support the use of ABI only as part of a comprehensive vascular assessment, with the use of other techniques to form a more complete clinical picture.’