Author’s response to reviews

Title: Evaluation of a simple tool to assess the results of Ponseti treatment for use by clubfoot therapists: a diagnostic accuracy study

Authors:
Tracey Smythe (tracey.smythe@lshtm.ac.uk)
Debra Mudariki (debramashoro@ymail.com)
Maxman Gova (gova64@icloud.com)
Allen Foster (allen.foster@lshtm.ac.uk)
Christopher Lavy (chris.lavy@ndorms.ox.ac.uk)

Version: 1 Date: 17 Jan 2019

Author’s response to reviews:

Please see attached cover letter files for response to Editor and reviewers. The unformatted response is included below:

JFAR-D-18-00176
Evaluation of a simple tool to assess the results of Ponseti treatment for use by clubfoot therapists: a diagnostic accuracy study
Tracey Smythe; Debra Mudariki; Maxman Gova; Allen Foster; Christopher Lavy
Journal of Foot and Ankle Research

Dear Editor,

Thank you for considering our manuscript for publication. Please find attached a manuscript version with 'tracked changes' of all changes made when revising the manuscript. In addition, we have:

1. Removed the job roles from the title page and only included full names, institutional addresses and email addresses for all authors.

2. Removed non-standard abbreviations (i.e: FAB, ICC, QoL and STARD) when only used once or twice within the manuscript. CTEV was replaced by ‘clubfoot’ in the manuscript to assist with ease of reading.

3. Removed secondary headings in discussion
4. Aligned some keywords to MeSH terms i.e. removed CTEV, changed measurement to measure

5. Reviewed the tense within article to ensure past tense is consistently used

6. Amalgamated the author contributions where two or more authors undertook a similar task.

With kind regards,

Tracey

JFAR-D-18-00176
Evaluation of a simple tool to assess the results of Ponseti treatment for use by clubfoot therapists: a diagnostic accuracy study
Tracey Smythe; Debra Mudariki; Maxman Gova; Allen Foster; Christopher Lavy
Journal of Foot and Ankle Research

Reviewer reports:
Reviewer #1 (Dr Kelly Gray): I would like to commend the authors on this piece of work. There is a clinical need for a tool which assists in identifying those with relapsing clubfoot. The main challenge with this work is that it identifies only those with significant relapse and may fail to recognise those with a good outcome who have early signs of relapse (which may then be able to be more easily managed). Having said that, I do feel this is a good simple tool which can very well serve the purpose that you have proposed it for. My suggestions therefore I feel fall under minor revisions. That is, with clear acknowledgment of the limitations this remains a useful tool. I wish the authors all the best in publication.

Thank you for your comments and suggestions. We reply in blue and underline the changes and revisions that we have made.

Minor revisions:

Page 5 - opening paragraph (Background) - In the final sentence, it would be useful to quantify the ratio of male to female risk or presentation of clubfoot.

Thank you for your recommendation. We have included the following information at the end of the final sentence of the opening paragraph:
Male sex is consistently associated with an increased risk of clubfoot; clubfoot affects twice as many boys as girls [3].

- 2nd paragraph, final line - consider changing 'Recurrence of the elements of the deformity is therefore less common' to 'recurrant of the elements of the deformity are therefore less common..'
Thank you for your suggestion. We have changed the wording in the final line of the second paragraph to:
Recurrent elements of the deformity are therefore less common after the child is four years old as growth of the foot decreases in speed.

- 3rd paragraph. There is reference that Ponseti-laaveg and Dimeglio are complicated to use. I don't entirely agree with this comment. In particular we find Dimeglio very easy to use. However, the flaw with both is that they are not validated to identify those who have relapse which require intervention.

We have clarified the wording with reference to the Ponseti-laaveg and Dimeglio Several scoring systems have been described for clubfoot; these include the Ponseti-Laaveg classification [9] and the Dimeglio classification [10], which are not validated to identify children with recurrent clubfoot who require intervention

- 3rd Paragraph - Consider removing the part of the sentence which states 'because it does not reflect the extent of the deformity'.

We have removed the suggested part of the sentence and the wording now reads:
The Pirani score [11] is frequently used to assess success during the corrective phase of treatment, however it is not validated for use in children of walking age.

- 3rd paragraph - final sentence - states' there is no agreed assessment of severity of recurrence' - to improve clarity, consider stating 'there remains no consensus on when intervention of recurrence should occur'

Thank you for your recommendation. We have changed the sentence to improve clarity:
There remains no consensus on when intervention of recurrence should occur and elements of the deformity that recur are typically noted under clinical examination and observation of function.

Page 6- the final sentence before the METHODS - consider removing 'including parent reported outcome measures about the key quality of life factors that affect the child and parent.'

We have removed the suggested wording and the sentence now reads:
To address this gap, we aimed to develop a user friendly, comprehensive tool to assess children of walking age who have undergone Ponseti treatment for clubfoot.

Page 10 (table 2) - I noted that all baseline Pirani scores were less than 4. There is a general agreement in the literature that a score of 5 or more corresponds to a more severe clubfoot which is also more likely to require an achilles tenotomy. It would be worth noting in the discussion section that your cohort are potentially less severe at baseline.

Thank you for your suggestion. We have included this information in the discussion section under study limitations, paragraph 7:
There were also study limitations. The cohort were potentially less severe at baseline, with mean initial Pirani scores of 4 or less. The results in those followed up are likely to be better than those for the cohort as they attended clinic appointments for longer and length of follow up is a predictor of good outcome.
Page 12 (Table 4) - I wondered for greater clarity, if this data may be presented in another way - for example as a bar graph. I found it visually challenging in its current form.

We have created bar graphs as a visual aid to assist with understanding of the data. We suggest that Table 4 is included as a web appendix for reader reference. As the data were complicated to display in one graph, we have included bar graphs for each questions answered.

Page 12 - consider presenting likelihood ratios - which may provide greater clinical relevance to clinicians.

We present likelihood ratios for the ACT score on page 13, one for positive and one for negative test results, under the heading of sensitivity and specificity. Table 3 and 4 provide that data that we use to calculate the likelihood ratios.

Page 16 Limitations - consider the following limitations - this does not identify children who require Tibialis anterior tendon transfer until they present with significant relapse. The score is excellent at providing a cross sectional assessment however further research is required to determine if it can detect earlier relapse (via the parent reported measures). For example some Tibialis Anterior Tendon transfer studies have noted that children who require TATT's often present with pain without other signs of structural relapse.

Thank you for your suggestion. We have included the following information in paragraph 8 of the discussion:
As non-specialist health workers regularly manage the treatment of clubfoot in low resource settings there is a need to provide appropriate tools to allow measurement and evaluation of their treatments. Further work is required to evaluate the ACT tool in other situations and with other cadres of clubfoot therapists. Also further exploration to differentiate children who score 9 or 10 with a good outcome from those who need referral is warranted; in particular the tool is not sensitive in identifying children who have a curvature in the front of the foot but who score high due to parent satisfaction, good footwear use and absence of pain. For example, the tool does not identify children who require tibialis anterior tendon transfer until they present with significant recurrent deformity. The score is excellent at providing a cross sectional assessment however further research is required to determine if it can detect earlier relapse via the parent reported measures.

I would consider removing the recommendation for yearly use of the tool. This should be done at every visit.

Thank you for your suggestion. The tool was used to assess children of walking age and has not been tested in children who have started casting within weeks of birth, and therefore completed casting and commenced bracing at approximately 2-3months of age. We believe that a yearly interval will allow for more accuracy when answering the questions regarding pain and use of shoes as the children will likely be walking.

Overall this is a good paper which is written well. I wish the authors all the best in it's journey to publication.

We thank you for your review.
Reviewer #2: Angela Evans (PhD)
Thank you for the opportunity to review this manuscript.
The development of a simple tool for clubfoot assessment post treatment that is quick, low-cost, repeatable and relevant is indeed helpful.
The development of the ACT tool is pragmatic and purposeful, and importantly includes parent rating as well as foot features and shoe use.

Overall the paper is well thought through and the study was well designed.

Thank you for your comments and suggestions. We reply in blue and underline the changes and revisions that we have made.

However, there is need for greater clarity regarding the 31% follow up rate and implications - this must be transparent and included in the abstract.

Thank you for your recommendation. We have included this information in the abstract, in both the method and results:

Method:
A literature review and a Delphi process were used to develop the Assessing Clubfoot Treatment (ACT) tool and score. We followed up children with clubfoot from a cohort treated between 2011 and 2013, in 2017.

Results:
The follow up rate was 31.2% (68 children). The ACT tool consisted of 4 questions;

Similarly the development was only based on opinion from 35 Ponseti trainers in Africa (the limitations of Delphi)

Thank you for your recommendation. We have included the following text in the Method of the abstract:

Method:
A literature review and a Delphi process based on the opinions of 35 Ponseti trainers in Africa were used to develop the Assessing Clubfoot Treatment (ACT) tool and score.

Likewise, these effects on the external validity and hence the *preliminary and limited findings of this study* need to be more effectively communicated.
I do not see this as preclusive of publication at all; on the contrary, I think this paper should be published but must make very clear its limitations and drive further work in this area.
Thank you for your suggestion. We have included the following information in paragraph 8 of the discussion to give an example of the limitations of the questions that the tool uses:

As non-specialist health workers regularly manage the treatment of clubfoot in low resource settings there is a need to provide appropriate tools to allow measurement and evaluation of their treatments. Further work is required to evaluate the ACT tool in other situations and with other cadres of clubfoot therapists. Also further exploration to differentiate children who score 9 or 10 with a good outcome from those who need referral is warranted; in particular the tool is not sensitive in identifying children who have a curvature in the front of the foot but who score high due to parent satisfaction, good footwear use and absence of pain. For example, the tool does not identify children who require tibialis anterior tendon transfer until they present with significant recurrent deformity. The score is excellent at
providing a cross sectional assessment however further research is required to determine if it can detect earlier relapse via the parent reported measures.

The ACT tool looks promising from all perspectives, but this paper needs to be very frank about the low follow up rate, one country setting, and implications of not including 69% of the initial cohort.

Thank you for your recommendation. We have included further clarification in the limitations section of the discussion (paragraph 7):
There were also study limitations. The study was undertaken in one clinic setting. Only 31.2% of the initial cohort were followed up. The cohort were potentially less severe at baseline, with mean initial Pirani scores of 4 or less. The results in those followed up (31.2%) are likely to be better than those for the cohort as they attended clinic appointments for longer and length of follow up is a predictor of good outcome. The tool is limited to one clinical examination, which restricts identification of pathology that is reliant on complex investigations. It is possible that results from the ACT tool may have influenced the decision to refer. Administering the tool first, but calculating the total score after the full clinical assessment, in addition to requiring agreement on the referral decision, should have reduced this potential for observer bias.

I am not sure about the PedsQL or healthcare satisfaction results as additional Files 3 and 4 were not included in those I received?
I will be happy to receive these an comment further.
Thank you.

Table 3: please correct recruitment numbers, 83+107+20 = 210; 8 less than 218 total, so not 100%.
Missing cases must be regarded as failed
- error in 3rd column from right - 50&

Thank you. We have corrected these errors.

The missing cases are regarded as failed and the number of children that did not finish casting is therefore 28. 83+107+28 = 218 and the % have been adjusted accordingly. We have also corrected the corresponding Figure (figure 1) with the number of enrolled children as n=218 and corrected the relevant percentages.

We corrected 50& to 50% in table 3.

Please temper results re sensitivity/specificity, with comment such as: 'given the small sample...'
Table 6 is important but needs to be clearer, and the use of % given many n<10 results is a little misleading.

We have included the suggested wording at the beginning of the paragraph with the results from the sensitivity and specificity tests:

Sensitivity and specificity
Given the small sample, 24 children required further intervention of which 19 scored 8 or less on the ACT score (sensitivity: 79%) and the remaining children scored 9 or 10.

We have adjusted the column size/width and row height to improve readability of Table 6. We have
maintained the use of % to align with the AMA Manual of style. We have included the Total (n) in the second row to provide reference for the denominators.

May I suggest amending the title:
‘Preliminary evaluation of a simple tool for the assessment of Ponseti treatment outcomes by non-medical personnel in low income settings’

Thank you for your suggestion. The STARD statement (Standards for Reporting of Diagnostic Accuracy Studies) that we have followed in the reporting this study was developed to improve the completeness and transparency of reports of diagnostic accuracy studies. The guidelines require that the study is identified as a study of diagnostic accuracy. We believe that the inclusion of “diagnostic accuracy study” in the title is therefore necessary. We use the terminology of ‘clubfoot therapist’ as this includes nursing staff, physiotherapists, community workers and rehabilitation technicians, and may encompass a broad spectrum of health care personnel. We have include ‘low income settings’ in the search terms to avoid the title being unwieldy and would therefore appreciate maintaining the title.

- and then please review and slightly re-frame the discussion to balance both the importance of the tool, and the qualified findings.

We have re-framed the discussion and removed the subheadings. To balance the importance of the tool and the qualified findings, we have included more detail on the study limitations in paragraph 7 and examples in paragraph 8 to make clear the research limitations and drive further work in this area.

There were also study limitations. The study was undertaken in one clinic setting. Only 31.2% of the initial cohort were followed up. The cohort were potentially less severe at baseline, with mean initial Pirani scores of 4 or less. The results in those followed up (31.2%) are likely to be better than those for the cohort as they attended clinic appointments for longer and length of follow up is a predictor of good outcome. The tool is limited to one clinical examination, which restricts identification of pathology that is reliant on complex investigations. It is possible that results from the ACT tool may have influenced the decision to refer. Administering the tool first, but calculating the total score after the full clinical assessment, in addition to requiring agreement on the referral decision, should have reduced this potential for observer bias.

As non-specialist health workers regularly manage the treatment of clubfoot in low resource settings there is a need to provide appropriate tools to allow measurement and evaluation of their treatments. Further work is required to evaluate the ACT tool in other situations and with other cadres of clubfoot therapists. Also further exploration to differentiate children who score 9 or 10 with a good outcome from those who need referral is warranted; in particular the tool is not sensitive in identifying children who have a curvature in the front of the foot but who score high due to parent satisfaction, good footwear use and absence of pain. For example, the tool does not identify children who require tibialis anterior tendon transfer until they present with significant recurrent deformity. The score is excellent at providing a cross sectional assessment however further research is required to determine if it can detect earlier relapse via the parent reported measures.

I am happy to review again, or comment further.

Well done with this important initiative.

Thank you for review.