Author’s response to reviews

Title: Measurement properties of instruments to assess pain in children and adolescents with cancer: A systematic review protocol

Authors:

Erik Loeffen (eah.loeffen@umcg.nl)
Jennifer N Stinson (jennifer.stinson@sickkids.ca)
Kathryn A Birnie (kathryn.birnie@sickkids.ca)
Monique van Dijk (m.vandijk.3@erasusmc.nl)
Ketan Kulkarni (Ketan.Kulkarni@iwk.nshealth.ca)
Mienke Rijsdijk (M.Rijsdijk-2@umcutrecht.nl)
Anna Font-Gonzalez (a.fontgonzalez@amc.uva.nl)
L Lee Dupuis (lee.dupuis@sickkids.ca)
Elvira C van Dalen (e.c.vandalen@amc.uva.nl)
Renée L Mulder (r.l.mulder@amc.uva.nl)
Fiona Campbell (fiona.campbell@sickkids.ca)
Wim JE Tissing (w.j.e.tissing@umcg.nl)
Marianne D van de Wetering (m.d.vandewetering@amc.uva.nl)
Faith Gibson (faith.gibson@gosh.nhs.uk)

Version: 1 Date: 07 Nov 2017

Author’s response to reviews:

NB. See also the attached cover letter for these answers (formatted to be read more easily), thank you.

First, we would to thank the editors for arranging the peer review process.
REVIEWERS COMMENTS AND RESPONSES

Reviewers

REVIEWER 1

Comment R1.1
I was very happy to read such a great article. This is a very important issue to be addressed as a systematic review and I look forward to the results.

Response R1.1
We thank reviewer 1 for his/her kind words and the helpful remarks.

Comment R1.2
I have small comments regarding the protocol. First I think that the explanation on how the age will be separate in the analysis could be more clear.

Response R1.2
We have added the following sentences to the ‘methods – synthesis of results’ section: “In addition an outcome table will be prepared which categorizes instruments according to age group, i.e. 0-0.99 years (infants), 1-2.99 years (children not able to speak or read), 3-6.99 years (children able to speak, not able to read), 7-12.99 years (children able to speak and read), 13-18 years (teenagers) . If the identified evidence leads to the conclusion that differently defined age groups are more appropriate, we will reconsider the definition of the age groups.” (lines 268-274).

Comment R1.3
Also if there will be a subgroup analysis (i.e. by age or gender) of the pain in those groups this should be stated. All best.

Response R1.3
We do not plan subgroup analyses for specific subgroups. However from the quality matrix (see response R3.9) and the outcome table that presents measurement instruments per age group, readers can per age group identify the measurement instrument with the strongest psychometric properties.

REVIEWER 2

Comment R2.1

I enjoyed reading the manuscript for your systematic review protocol, which follows guidelines for systematic review methods very well, and clearly has important implications for the field. I only have some minor comments:

Response R2.1

We thank reviewer 2 for his/her kind words and the helpful remarks.

Comment R2.2

You have planned what appears to be a very robust and considered search strategy, but I notice the lack of attention to grey literature, for example theses/dissertations. This may be worthwhile considering as an addition to your database searches as it may reduce the impact of any publication bias. I notice also that you plan to conduct backwards citation chasing (checking references of your included studies) but not forwards citation chasing (checking the reference lists of papers that have since cited your included studies). This could also be a worthwhile addition to your search strategy, especially since this method focuses on more recent literature that may have been missed.

Response R2.2

Thank you for these suggestions. Regarding grey literature, we have added that we will hand search conference proceedings and trial registers to identify promising ongoing and recently completed studies. See line 187-196: “In addition, to identify promising ongoing […] Clinical Trials Registry Platform (ICTRP).”.

We deliberately chose not to include theses/dissertations, because these publications are not always peer-reviewed (and thus appraised for basic scientific quality). In addition we are confident that quality research from theses/dissertations is usually published in peer-reviewed journals and will therefore not be missed. Regarding the forward citation chasing, we think this
is a very worthwhile addition and have added this to the protocol, see lines 184-186: “[..] we will check all references of included studies (backwards citation chasing) and all studies that referenced the included studies (forward citation chasing).”.

Comment R2.3

Throughout your paper you refer to instruments to assess pain in children with cancer, but in your inclusion criteria you state that you will include studies where the whole sample is <25 years or where the mean/median age is <16, meaning that study samples could include adolescents or even young adults. I wonder if the phrase "children and young people" might be more appropriate to use throughout your paper?

Response R2.3

The definition for this inclusion criterion has come forward from previous systematic reviews that we have performed in this field. In these reviews, we defined studies with ‘children’ as studies that included only patients ≤18 years of age. However, this would lead to the exclusion of studies that included one or only a few patients aged above 18 years, which in our eyes lead to the loss of valuable data. In addition we learned that not all studies provide proper descriptives (age range, mean or median ages). Therefore we have chosen for the current definition, in which we focus on children and accept the occasional included young adult. However we agree that changing the phrasing might be more appropriate, and we think that in our case “children and adolescents” describes the population best. We have adjusted the manuscript accordingly, thank you.

Comment R2.4

On line 226 you state that you will seek to identify cues that imply publication bias. This section could be strengthened by describing how you intend to identify those cues and, indeed, what those cues might be.

Response R2.4

Thank you for this comment. For reviews that evaluate clinimetric properties there is no established method to evaluate publication bias (e.g. as funnel plots are in meta-analyses). Therefore we are bound to commenting on cues for publication bias in a qualitative manner. For instance, a possible cue might be that we only find studies evaluating measurement tools that
prove to have strong clinimetric properties. This might mean that studies evaluating measurement tools that have less strong clinimetric properties remain unpublished. We have added the following text: “[..] that imply this bias, e.g. identification of merely studies that report measurement instruments with strong clinimetric properties.” (lines 257-259) and “Also there is no established methodology for this, e.g. as there is for a meta-analysis with a funnel plot. Therefore we will comment upon this in a qualitative manner.” (lines 261-263).

Comment R2.5

In the discussion you describe plans to produce tables of characteristics and outcomes of included studies. Presumably you will also have a narrative synthesis to bring together the results of individual studies? This should be highlighted in the protocol.

Response R2.5

Indeed we will, we have added this to the appropriate section, see line 280: “In addition, we will present a narrative synthesis discussing our findings.”.

Comment R2.6

On line 239 you describe part of the rationale for this review is that there are clear differences between adults and children, but in the introduction you do not discuss the adult literature. The paper could be strengthened by making brief reference to the adult literature and in particular whether any systematic reviews have been conducted in adult populations.

Response R2.6

This is a very valuable suggestion, for which we thank reviewer 2. We have added a paragraph to the introduction, see lines 110-117: “In previous research, focus has been put on synthesizing knowledge regarding pain measurement tools in other populations. A systematic review on pain intensity assessment tools for adults concluded that numerical rating scales are applicable for unidimensional assessment of pain intensity in these patients.(10) Another recent systematic review focused on non-responsive adults, and found that there are multiple observational pain assessment tools available for this population.(11) However, there are clear differences between children and adults, and there are various types of pain that children with cancer experience (e.g. neuropathic pain) that require specific guidance.”.
Comment R3.1

The aim of the review is to inform recommendations for health professionals on the assessment and treatment of pain in children with cancer which is laudable. A little more structure on the approach to synthesising and ordering the findings will benefit the review aims.

Response R3.1

We thank reviewer 3 for his/her kind words and the helpful remarks.

Searches

Comment R3.2

I'd like to see a bit more on the search strategy in the main body text about the keywords used to build the strategy and date limitations.

Response R3.2

We have added example search terms (both free text terms and MeSH headings) in the text, see lines 171-178. The full search strategy is stated in Table 2, thus this can also be easily consulted by future readers of the full article. As stated in line 169, we do not pose limitations regarding publication date.

Methods

Comment R3.3

Some info on the stages of screening titles and abstracts is unnecessarily detailed. Some of the phrasing a bit clumsy such as Line 185 "Reviewers will be instructed to be biased towards inclusion". Maybe amend to "a conservative approach to inclusion will be employed to ensure relevant titles are not excluded".

Response R3.3

Thank you for these suggestions. We have shortened the section on title selection and we have amended the line as suggested by the reviewers, which we agree is a more appropriate and better put phrasing.
Comment R3.4

Are two reviewers screening all titles and abstracts or will the second reviewer only screen a proportion of the first reviewer's? If so, what proportion?

Response R3.4

In both screening phases the first and second reviewer will review all titles and abstracts. To make this more clear, we have rephrased both line 207 and line 217 to “The entire selection will be performed in duplicate by two independent reviewers ([..])”.

Comment R3.5

At abstract selection stage, the position of second reviewer is divided among seven people. What is the purpose/value of this and how will it affect consistency of study selection?

Response R3.5

In the guideline project we are working on, all working groups consist of multiple members. To familiarize all members with the process of study selection (so that all members know how we performed the actual review) and to divide the workload, we chose to divide the second reviewer’s work among the working group members. To not compromise consistency, the first reviewer is one person (and the same from the title selection phase), and we will examine agreement between reviewer 1 and all individual reviewer 2 portions.

Comment R3.6

Some info is erroneous such as "citations classified as 'include' by at least one reviewer will be included." This goes without saying. More detail about the eligibility criteria, particularly on the types of studies would be useful. For example, will this review include only quantitative studies (RCTs, single arm, case, series, cohort, observational) of pain measurement instruments in children or will qualitative studies examining user or health professionals perceptions and reflections of these tools also be integrated (mixed methods review).

Response R3.6

Thank you for this comment. Perhaps we were not entirely clear, but we think the information in the cited sentence is not erroneous. Normally in systematic review selection procedures discrepancies are resolved by consensus, thus a citation that is initially marked as ‘include’ by one of the reviewers can still be excluded after discussion. In our title selection procedure, we
will not have such a discussion and we will just include all citations of which at least one reviewer thought it should be included. This is part of the aforementioned bias towards inclusion / conservative approach. To clarify this in the text, we have added: “All citations classified as ‘include’ by at least one reviewer will be included, even if the second reviewer classified the citation as ‘exclude’ (no discussion will be held).” (lines 211-212). We agree with reviewer 3 that we should have stated the type of studies more clear (limited to quantitative studies), we have added this to line 160.

Data extraction

Comment R3.7

There is minimal information on precisely what data will be extracted from the primary studies of instruments included and how those data will be synthesised. For example, it is not clear whether this review will simply outline which tools are out there (mapping review), or how it aims to rank them in terms of clinimetric utility? The prospero registration mentions that a data abstraction form has been developed, maybe include this as an appendix?

Response R3.7

Regarding the data that will be extracted from the included studies, see lines 230 to 236. In addition, we agree with Reviewer 3 that the addition of the data extraction form is a valuable addition, thus we have added this as supplementary data. We also agree with Reviewer 3 that the synthesis of evidence section can profit from elaboration. Therefore, we have added the following (lines 268-280): “In addition an outcome table […] narrative synthesis discussing our findings.”.

Comment R3.8

Moreover the review questions in Table 1 need elaboration. What constitutes "good clinimetric properties" and what, for example, would constitute not good? There seems to be a bit of a leap between understanding how studies will be selected into the review, to how they will be appraised. I understand that this process is subject to development throughout the research itself but if the protocol is to be published as an academic publication in itself, it needs to have a clearer direction about what data is anticipated and how it will be synthesised in order to serve as something more empirically sound than the CRD protocol registration.
Response R3.8

We have changed the word “good” in Table 1 to “strong” to have the phrasing in line with the main text. In formulating the research questions, we initially included all outcomes in the question. This however lead to long questions that were hard to comprehend. Therefore we chose to phrase it as “strong clinimetric properties”, which means measurement instruments with proven validity, reliability, etc. Regarding anticipated data, we have added several parts of texts to the “quality assessment” and “synthesis of results” sections, see also Response R3.9.

Quality Assessment

Comment R3.9

The quality assessment section details checklists that will be used and explains that criteria by Cohen et al will be used to categorise the instruments but it’s not clear whether the 3 categories 'well-established assessment', 'approaching well-established assessment', 'promising assessment' would provide a meaningful hierarchy for readers who are interested in learning which instruments were found to be most, or indeed, least promising for assessing pain in children with cancer from the systematic review. What sorts of criteria, other than methodological quality, would determine which instruments would rank the highest, such as: prevalence in the literature, most extensive validation or service user valuation. More details are required to inform readers what the authors regard as the most important or promising features of relevant measurement instruments.

Response R3.9

In incorporating the feedback from the reviewers and adjusting the protocol, we came to the insight that the added value of the Cohen classification (i.e. in addition to the COSMIN appraisals) is very minor. Therefore we chose to remove this from our protocol and review. To provide the reader with a quick and comprehensive overview of our findings (and thus identify the most appropriate measurement instrument for their purpose/paperation), we have decided to develop a quality matrix that will incorporate this information and the quality of evidence scores. We have added the following: “To provide a comprehensive overview, we will develop a quality matrix. Herein we will list all included measurement instruments and the identified information on purpose (self- or observer-report, type of pain), number of studies, population / age group, and per outcome COSMIN quality score (four-level). This will allow the reader to quickly identify the appropriate measurement instrument with the strongest clinimetric properties.” (lines 275-279).
Discussion

Comment R3.10

In the discussion the authors highlight a number of important considerations that would determine how instruments could be categorised in the review, such as developmental stages across ages and specific types of pain related to their disease. These seem like important factors that should dictate the way the studies are grouped in the synthesis. The authors acknowledge the impact that recognition and treatment of pain in this group has on their quality of life. I would like to see more discussion of the issues relevant of communication, particular for young, pre-language children and the potential impact of self versus parent/clinician proxy measurement of pain. In many cases children are dependent on their carers', health professionals taking action on their behalf and therefore the use of instruments in these cases should be acknowledged.

Response R3.10

We thank Reviewer 3 for these comments. Regarding synthesis of the evidence, see answers R3.8 and R3.9. Regarding the comments on the discussion of our manuscript, we acknowledge that these are factors that should be touched upon when placing the findings into context. We believe however this is best done in the discussion of the final review, when these matters can be supplemented with the knowledge generated in the present study. For the discussion of this protocol, we think it is mainly important to touch upon why this review is needed. Nevertheless we agree with Reviewer 3 that it is pivotal to underline the importance of proxy measurement instruments. We have added to the discussion (lines 291-295): “Although self-report is often regarded as the ideal method for assessing pain, in young children this is not always possible. Therefore, these patients have to rely on the people around them (family but also healthcare professionals) to notice and act upon their pain. This underlines the importance of having clinimetrically sound measurement instrument not just for self-report, but also for proxy report.”.

Comment R3.11

Table 1: not all colours are added as described.

Response R3.11

This is because there were not outcomes with a score of 1-3. We have changed “highlighted orange” to “none in this case”.
Comment R3.12

Table 2: I would have liked to have seen the numbers generated by the sample search strategy in order to gauge appropriateness of the search.

Response R3.12

Thank you for this comment. We have added the numbers generated by the search as performed with the stated strategy on March 22nd 2017. See Table 2.

Comment R3.13

Line 145: Define VAS on first use and in List of Abbreviations. Also check all abbrevs are added to this list such as COSMIN.

Response R3.13

We have checked the manuscript for definitions and use of abbreviations, and made the List of Abbreviations complete. It should be noted that COMFORT (the measurement tool name) is neither added to this list nor defined as this is a name and not an abbreviation (although it seems like one).

We like to thank all reviewers again for their time and suggestions.