PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	International Mixed Methods Study Protocol to Develop a Patient-Reported Outcome Measure for all Types of Chronic Wounds (the WOUND-Q)
AUTHORS	Klassen, Anne; van Haren, Emiel; Cross, Karen; Fan, Kenneth L; Gibbons, Chris; Hoogbergen, Maarten M; Longmire, Natasha M; Poulsen, Lotte; Sorensen, Jens Ahm; Squitieri, Lee; Tsangaris, Elena; van Alphen, Tert C; van Dishoeck, Anne-Margreet; Vasilic, Dali; Pusic, Andrea L

VERSION 1 – REVIEW

REVIEWER	Agata Janowska Department of Dermatology, University of Pisa, Italy
REVIEW RETURNED	26-Jul-2019

GENERAL COMMENTS	WOUND-Q is a interesting and ambitious project

REVIEWER	Professor Hayley Hutchings Swansea University, U.K.
REVIEW RETURNED	04-Nov-2019

GENERAL COMMENTS	On the whole this is a well written manuscript of a proposed study to develop a generic PROM for wounds. As this is a protocol for a proposed study I feel it would be better written in the future tense throughout the manuscript. Currently it reads in the present tense which reads like the bulk work has already been completed. I would have liked to see stronger justification for why the WOUND-Q is needed when there is Wound-QoL already in existence for use across multiple wound types. Why was a new tool deemed necessary rather than undertaking modern psychometric methods with this existing PROM? I thought the patient and public involvement section was a little weak. Patients appear to be involved as participants but not as part of the wider research team. There has been excellent guidance recently published on patient and public involvement and I believe the authors could strengthen the public and patient involvement in the study.

REVIEWER	Christine Blome, PhD University Medical Center Hamburg-Eppendorf, Institute for Health Services Research in Dermatology and Nursing, Germany

	I am one of the authors of the Patient-Reported Outcomes Measure "Wound-QoL".
REVIEW RETURNED	11-Nov-2019

GENERAL COMMENTS	bmjopen-2019-032332 Review by Christine Blome This manuscript describes the study protocol for the development (and, as an outlook, the validation) of a new patient-reported outcomes measure (PROM) for use in people with chronic wounds. The rationale for such a development in spite of several existing instrument is clearly described in the Introduction, i.e. no PROM for all types of chronic wounds exists that has been developed rigorously using both classic and modern psychometric methods (it might be added that most existing PROMs focus on HRQoL only). Especially for a PROM development study, it makes sense to publish the research protocol in order to ensure that (a) any protocol deviations will have to be made transparent in the latter publication of the study and (b) the research community will be informed about this upcoming PROM development, thereby potentially reducing the waste of research resources. The manuscript is very well-written, clear, and comprehensive. For example, there is a highly specific description of the statistics to be used in the item selection process. The manuscript might however benefit from some changes and amendments as listed below. Strengths and limitations section: Here, only strengths are listed; at least one of the limitations discussed in the Discussion section should be mentioned. Strengths and limitations section & general: You write that an “internationally-applicable” PROM will be developed, which is true, but only patients / clinicians from highly developed Western countries will be included. Applicability of the PROM to patients from different cultures might be discussed as a limitation. Page 6, line 17-21: An explanation or reference on why total scores are problematic in clinical trials would be helpful. In addition, it may not be generally true that total scores are hard to interpret, especially if the examples given in brackets (varying directions and sizes of scores) are not pertinent. The PROM to be developed shall include different scales that measure different concepts of interest to patients and providers; these concepts will be defined in the qualitative study part. The introduction suggests that these concepts will only include HRQoL and symptoms. It would be helpful to be more explicit about this – shall the new PROM cover any concept that is highly important to patients/providers (e.g., burden of treatment; caregiver burden; etc.), or are concepts confined to symptoms and HRQoL? Page 7, line 35-42: In purposeful sampling, I believe it would also be important to include patients with different levels of education and/or health literacy in order to ensure comprehensibility of the PROM in different patient groups. Also, is the PROM meant to also be applicable to patients with mild cognitive impairment? This might be an important point as chronic wounds are especially prevalent in elderly people.

	Page 8, line 22: Why do you plan to use Excel instead of a software specific for qualitative analysis? Page 8, line 36-40: "Also, performing interviews and analysis at the same time for member-checking takes place to confirm that the COI identified in interviews is confirmed in subsequent interviews." In which way exactly and for which party of analysis shall member-checking be performed? I assume that each scale of the PROM shall be based on a reflective model instead of a formative model, as unidimensionality will be checked. This could be made more explicit. Page 10, line 26: Which experts will be recruited – physicians only? Or also nurses and other professional groups? Do you plan to recruit wound care experts who are specialized in different wound types, or wound care experts in general, assuming they do have the broad knowledge? Page 14, line 10-24: Patients take part in the study as participants only. Did you consider working with patient research partners, too? Otherwise, speaking of "engaging" patients might be misleading. Page 14, 17-22: It is entirely not clear to me why participation of the same patients in both qualitative interviews and cognitive interviews will be an advantage; in contrast, it might make more sense to ask patients without knowledge of the development process and the concepts to be measured to judge the draft PROM. Figure 2: My personal experience is that in PROM translation, it makes sense to also ask a professional translator to proof-read the final version, as even when working with professionals in the translation and back translation process, oftentimes still errors are uncovered in proof read step; you might consider doing so in addition to the proof-read by clinicians. Minor points: Page 5, line 35: "such as pain" instead of "such pain"; Page 5, line 50: reference #13 is not a review, which is why it should not say "[12-15]" but [12,14,15]"; Page 6, line 8: "Wound-QoL [27]" should say "Wound-QoL [28]"; Page 6, line 26: "all types of chronic wounds" instead of "…type…"; Page 9, line 8: Is there a word missing after "across"?; Page 10, line 38: "Table 2" should be "Figure 2".

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 Comment: WOUND-Q is an interesting and ambitious project.

Author response: Thank you.

Reviewer 2 Comment: On the whole this is a well written manuscript of a proposed study to develop a generic PROM for wounds. As this is a protocol for a proposed study I feel it would be better written in the future tense throughout the manuscript. Currently it reads in the present tense which reads like the bulk work has already been completed.

Author response: We can change this if the editor requests but have used this tense in other protocol papers.

Reviewer 2 Comment: I would have liked to see stronger justification for why the WOUND-Q is needed when there is Wound-QoL already in existence for use across multiple wound types. Why was a new tool deemed necessary rather than undertaking modern psychometric methods with this existing PROM?

Author response: We have added to the introduction to justify the creation of a new tool when the Wound-QoL already exists. Briefly, the Wound-QoL covers a range of important concepts in via 3 scales and a total score. For example, the body scale asks about pain, exudate and sleep impact. The WOUND-Q is more granular as it has separate independently functioning scales for each important concept.

Reviewer 2 Comment: I thought the patient and public involvement section was a little weak. Patients appear to be involved as participants but not as part of the wider research team. There has been excellent guidance recently published on patient and public involvement and I believe the authors could strengthen the public and patient involvement in the study.

Author response: We acknowledge that our approach to patient and public participation is limited. While we do not have patients as members of our research team, we involve a large sample of patients in the development of the WOUND-Q. We find that ongoing involvement of patients as participants in the qualitative and cognitive interviews is an effective means for ensuring the content of the scales resonate with patient and measure the outcomes they care about. We have clarified the role of wound care experts, who played a role in the scale development, by revising a sentence in the research team meeting section and the patient public involvement section.

Reviewer 3 Comment: This manuscript describes the study protocol for the development (and, as an outlook, the validation) of a new patient-reported outcome measure (PROM) for use in people with chronic wounds. The rationale for such a development in spite of several existing instrument is clearly described in the Introduction, i.e. no PROM for all types of chronic wounds exists that has been developed rigorously using both classic and modern psychometric methods (it might be added that most existing PROMs focus on HRQoL only). Especially for a PROM development study, it makes sense to publish the research protocol in order to ensure that (a) any protocol deviations will have to be made transparent in the latter publication of the study and (b) the research community will be informed about this upcoming PROM development, thereby potentially reducing the waste of research resources. The manuscript is very well-written, clear, and comprehensive. For example, there is a highly specific description of the statistics to be used in the item selection process. The manuscript might however benefit from some changes and amendments as listed below.

Strengths and limitations section: Here, only strengths are listed; at least one of the limitations discussed in the Discussion section should be mentioned.

Author response: We added the following limitations:

"A limitation of our study is that patient involvement does not include membership in the research team. Another limitation is that the WOUND-Q field-test takes place only in high income countries."

Reviewer 3 Comment: Strengths and limitations section & general: You write that an "internationally-applicable" PROM will be developed, which is true, but only patients / clinicians from highly developed Western countries will be included. Applicability of the PROM to patients from different cultures might be discussed as a limitation.

Author response: We changed this point to read as follows:

"Recruitment of an international sample makes it possible to develop a patient-reported outcome measure (PROM) that reflects the concerns of patients in multiple countries."

Reviewer 3 Comment: Page 6, line 17-21: An explanation or reference on why total scores are problematic in clinical trials would be helpful.

Author response: We have revised this section and created a new paragraph in the introduction to expand on limitations of total scores that add up subscales.

Reviewer 3 Comment: In addition, it may not be generally true that total scores are hard to interpret, especially if the examples given in brackets (varying directions and sizes of scores) are not pertinent.

Author response: When scales are added together to get a total score, especially if the item set for each scale covers multiple concepts, it is hard to know what the scores mean (the concept of interest). An alternative approach is to have a set of independently functioning scales that each measure a unidimensional concept of interest.

Reviewer 3 Comment: The PROM to be developed shall include different scales that measure different concepts of interest to patients and providers; these concepts will be defined in the qualitative study part. The introduction suggests that these concepts will only include HRQoL and symptoms. It would be helpful to be more explicit about this – shall the new PROM cover any concept that is highly important to patients/providers (e.g., burden of treatment; caregiver burden; etc.), or are concepts confined to symptoms and HRQoL?

Author response: We have added the following sentence to the introduction:

"The WOUND-Q will contain a comprehensive set of independently functioning scales designed to measure outcomes that matter to patients with any type of chronic wound, as well as scales to measure patients experience of wound care."

Reviewer 3 Comment: Page 7, line 35-42: In purposeful sampling, I believe it would also be important to include patients with different levels of education and/or health literacy in order to ensure of the PROM in different patient groups.

Author response: We agree with these points and seek to include as varied a sample as possible in the qualitative study. In forming the scales, we strive to ensure the grade reading level is as low as possible. In addition, we do extensive cognitive debriefing interviews to ensure that each scale's content is comprehensive, relevant and, importantly, comprehensible to all participants.

Reviewer 3 Comment: Also, is the PROM meant to also be applicable to patients with mild cognitive impairment? This might be an important point as chronic wounds are especially prevalent in elderly people.

Author response: We agree with this important point. WOUND-Q field-test scales vary in terms of their Flesch-Kincaid grade reading levels. Only 1 scale has a grade reading level >6. Of the remainder, 3 scales are grade 1, 3 scales are grade 3, 6 scales are grade 4, and 2 scales are grade 5. We will publish the grade reading level in the psychometric field-test study paper in order to guide their uptake for different demographics.

Reviewer 3 Comment: Page 8, line 22: Why do you plan to use Excel instead of a software specific for qualitative analysis?

Author response: We have used NVivo software in past studies to develop PROMs, e.g., the CLEFT-Q study [1]. We find that data analysis using Excel is better for PROM development. In Excel we use columns for the multiple levels of coding (domains, major and minor themes). We can then sort by these to perform constant comparison and to examine saturation.

1. Wong Riff KWY, et al. What matters to patients with cleft lip and/or palate: an international qualitative study informing the development of the CLEFT-Q. Cleft Palate Craniofac J 2018; 55(3):442-50.

Reviewer 3 Comment: Page 8, line 36-40: "Also, performing interviews and analysis at the same time for member-checking takes place to confirm that the COI identified in interviews is confirmed in subsequent interviews." In which way exactly and for which party of analysis shall member-checking be performed?

Author response: We have revised the paragraph on rigor in qualitative research and hope that this revision clarifies our approach.

Reviewer 3 Comment: I assume that each scale of the PROM shall be based on a reflective model instead of a formative model, as unidimensionality will be checked. This could be made more explicit.

Author response: The WOUND-Q is aligned with the reflective model. We have made this clearer in our Method section, which now reads:

"Scale development is informed by the Rasch Measurement Theory (RMT) approach [30, 45]. In this approach, a pool of items that are reflective of the underlying constructs are derived from the qualitative data to create, for each scale, a conformable set of items that together map out a construct on a clinical hierarchy."

Reviewer 3 Comment: Page 10, line 26: Which experts will be recruited – physicians only? Or also nurses and other professional groups? Do you plan to recruit wound care experts who are specialized in different wound types, or wound care experts in general, assuming they do have the broad knowledge?

Author response: Experts include plastic surgeons, vascular surgeons, general surgeon and nurse practitioners. All experts have extensive and broad expertise in the management of chronic wound care.

Reviewer 3 Comment: Page 14, line 10-24: Patients take part in the study as participants only. Did you consider working with patient research partners, too? Otherwise, speaking of "engaging" patients might be misleading.

Author response: The patients are participants rather than research partners. We have changed the word "engage" to "involve" in order to ensure the distinction is clear.

Reviewer 3 Comment: Page 14, 17-22: It is entirely not clear to me why participation of the same patients in both qualitative interviews and cognitive interviews will be an advantage; in contrast, it might make more sense to ask patients without knowledge of the development process and the concepts to be measured to judge the draft PROM.

Author response: From our experience, there are various reasons for using the same participants in both phases of the study. Such participants are almost always keen to continue to be involved and have a lot to offer. Such participants already understand the purpose of the study, and they are invested in making sure the PROM reflects the concerns of patients like them. Such participants can tell us if the scales resonate with their experiences discussed in the qualitative interview. For example, one participant said the following: "There were a couple times where I actually felt a little emotional because the question really hits the nail on the head. You seem to get it. Sometimes people that are in your life don't get it, so when you read a question that really hits home, it's nice. Someone actually gets it." On a practical point, use of the same participants reduces the number of patients that sites need to recruit.

Reviewer 3 Comment: Figure 2: My personal experience is that in PROM translation, it makes sense to also ask a professional translator to proof-read the final version, as even when working with professionals in the translation and back translation process, oftentimes still errors are uncovered in proof read step; you might consider doing so in addition to the proof-read by clinicians.

Author response: Thank you for this suggestion. Unfortunately, the translations are completed. The translations were shown to 38 participants in a series of cognitive debriefing interviews, which led to changes to 11 items and the instructions. A few errors and inconsistencies were spotted by the experts who proof-read the final version as the last step.

Reviewer 3 Comment: Page 5, line 35: "such as pain" instead of "such pain";

Author response: We have made this change

Reviewer 3 Comment: Page 5, line 50: reference #13 is not a review, which is why it should not say "[12-15]" but [12,14,15]"

Author response: The Gorecki citations (13 and 25) were mixed up and have been exchanged.

Reviewer 3 Comment:

Page 6, line 26: “all types of chronic wounds” instead of “…type…”;

Author response: We have made this change to the text and title.

Page 9, line 8: Is there a word missing after “across”?;

Author response: We removed the word “across”.

Reviewer 3 Comment: Page 10, line 38: “Table 2” should be “Figure 2”.

Author response: We have made this change.

VERSION 2 – REVIEW

REVIEWER	Professor Hayley Hutchings Swansea University, UK
REVIEW RETURNED	11-Dec-2019

GENERAL COMMENTS	I'm happy that the authors have largely addressed my previous review comments. The only exception is the patient and public involvement section. The fact that patients were not involved as part of the research team has been stated as a limitation in the strengths and weaknesses section, but there is no reference to it in the PPI section itself. I think it may also be worth referring to guidance regarding good practice for PPI (https://www.invo.org.uk/wp-content/uploads/2019/02/71110_A4_Public_Involvement_Standards_v4_WEB.pdf).

REVIEWER	Christine Blome University Medical Center Hamburg-Eppendorf, Institute for Health Services Research in Dermatology and Nursing, Germany I am one of the authors of the Patient-Reported Outcomes Measure "Wound-QoL".
REVIEW RETURNED	20-Dec-2019

GENERAL COMMENTS	(1) As has become clear from the reply-to-reviewers, the WOUND-Q has already been developed and translated, i.e. the study protocol has already been implemented. I think it would be helpful to make this transparent in the manuscript, as it also implies that it will not be possible to implement in the protocol any suggestions on changed procedures made by the peer reviewers. Has the study protocol been finalized in the version submitted to BMJ Open already before conducting the study, or were any changes made in the process of study conduction? (2) Thank you for the helpful elaboration of the advantages of including the same participants in both qualitative interviews and cognitive interviews. However, I believe that it is also important to ask participants “with a fresh pair of eyes” for their feedback on a newly-developed instrument, i.e. patients without knowledge of the concepts and the development process (which will also be the case in most participants who will be asked to complete the Wound-Q in future research). As the cognitive debriefing has already been completed, I believe this point should be discussed as a possible limitation. (I would like to add that my first review

	read quite impolitely “It is entirely not clear to me” – what I wanted to say was “It is not entirely clear to me” – my apologies for this typo!) (3) With regard to my question whether the purposeful sampling also included level of education and/or health literacy as a criterion: Thank you for the additional information. As the cognitive debriefing has already been completed, I understand that it is not possible to make this change to the protocol anymore.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1 Comment: I'm happy that the authors have largely addressed my previous review comments. The only exception is the patient and public involvement section. The fact that patients were not involved as part of the research team has been stated as a limitation in the strengths and weaknesses section, but there is no reference to it in the PPI section itself. I think it may also be worth referring to guidance regarding good practice for PPI (https://www.invo.org.uk/wp-/uploads/2019/02/71110_A4_Public_Involvement_Standards_v4_WEB.pdf).

Author response: We have added a sentence to the start of the **PPI section** to restate that a limitation of our study is the lack of patient involvement on the research team.

Reviewer 3 Comment: (1) As has become clear from the reply-to-reviewers, the WOUND-Q has already been developed and translated, i.e. the study protocol has already been implemented. I think it would be helpful to make this transparent in the manuscript, as it also implies that it will not be possible to implement in the protocol any suggestions on changed procedures made by the peer reviewers. Has the study protocol been finalized in the version submitted to BMJ Open already before conducting the study, or were any changes made in the process of study conduction?

Author response: We have added the following sentence to the start of the **Subsequent phases** section to indicate progress made on the study to date: “The phase II field-test study is currently ongoing and will be completed in 2020.” The study protocol has not changed other than to add additional field-test sites that ask if they can participate.

Reviewer 3 Comment: (2) Thank you for the helpful elaboration of the advantages of including the same participants in both qualitative interviews and cognitive interviews. However, I believe that it is also important to ask participants “with a fresh pair of eyes” for their feedback on a newly-developed instrument, i.e. patients without knowledge of the concepts and the development process (which will also be the case in most participants who will be asked to complete the Wound-Q in future research). As the cognitive debriefing has already been completed, I believe this point should be discussed as a possible limitation. (I would like to add that my first review read quite impolitely “It is entirely not clear to me” – what I wanted to say was “It is not entirely clear to me” – my apologies for this typo!)

Author response: We added the following sentence to the **PPI section** to address: “We recognize a limitation of using the same participants twice could be that participants not involved in the initial phase may provide new insights.”

Reviewer 3 Comment: (3) With regard to my question whether the purposeful sampling also included level of education and/or health literacy as a criterion: Thank you for the additional information. As the cognitive debriefing has already been completed, I understand that it is not possible to make this change to the protocol anymore.

Author response: It is correct that changes are no longer possible at this time. However, our experience is that items that are challenging for people to understand (eg, ambiguous, unclear, not easy to translate) will exhibit poor item fit to the Rasch model. Poorly fitting items will be candidate items to delete from the scale during the psychometric analysis.