PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Improving the patient-centred care of children with life-altering skin conditions using feedback from electronic patient-reported outcome measures: Protocol for a hybrid effectiveness-implementation study (PEDS-ePROM)
AUTHORS	Tyack, Zephanie; Simons, Megan; McPhail, Steven; Harvey, Gillian; Zappala, Tania; Ware, Robert; Kimble, Roy

VERSION 1 – REVIEW

REVIEWER	Christina Zigler Duke University School of Medicine USA
REVIEW RETURNED	28-Jul-2020

GENERAL COMMENTS	The study aim is to investigate: (1) the effectiveness of a patient-centred care intervention using feedback from electronic patient-reported outcome measures (PROMs; PEDS-ePROM intervention) on health outcomes, referrals, and treatment satisfaction; and (2) the implementation of PEDS-ePROM by assessing acceptability and sustainability of the intervention and study processes. The first aim looks to see if the intervention improved health-related quality of life (HRQoL; primary outcome), scores on other PROMs, referrals, and overall tx satisfaction. The second aim looks at the overall implementation process of the intervention. From the perspective of the child and caregiver, they will explore acceptability and sustainability. There are a number of strengths of the study. They utilize a pragmatic randomized-controlled trial (varying presence of the PROM graphical display) and mixed methods with quantitative and qualitative data being collected from multiple sources (child, caregiver, EHR, clinicians). They also include a batter of PROMs, including generic scales and disease-specific measures. There were a number of places where the protocol write-up could be streamlined and some of the choices better justified. I have include specific questions below: 1. Hypotheses 2: The authors mention the PEDS-ePROM and ePROM but both of these terms were not fully defined until later in the paper. 2. It is not clear if the authors considered the impact of combining proxy and self-report into the same sample. It has been shown that agreement between children and caregivers is often low on PROMs. More justification for this choice should be provided. 3. It was not clear if randomization would be done within each target population (burn scars vs haemangiomas). What do the authors expect to be the final breakdown of the sample? (ex. 50%

	burns and 50% haemangiomas?) This has implications for the secondary data analysis. a. Do the authors expect severity of HRQoL to be different among the two groups of patients? b. Would the authors expect the intervention to be better received by one group or the other? 4. It was not clear why “referrals” were chosen as an outcome. Is a higher number of referrals a positive or a negative? 5. It was not clear why treatment satisfaction was chosen as an outcome. 6. Are the authors planning to stratify randomization among age groups? We know that age makes a difference in HRQoL reporting and again there are implications for combining proxy vs self-report. It would strengthen the paper if the authors could provide justification around their choice. a. Inclusion criteria – a lower bound for age is not specified. Are the authors planning to recruit infants? 7. Why does the study team believe that the graphical display will significantly improve outcomes over just receiving the intervention? 8. More information is needed around the power calculation and sample size justification. a. How was the effect size chosen? What metric is the effect size reported in? Is this for the primary outcome (PedsQL)? b. No attrition or loss of follow-up was taken into account. What happens if a number of people drop out? 9. Will the study team plan to take into account intervention ‘dosage’? (E.g. potential differences if child gets PROM feedback 6 times over 6 months versus 2 times). 10. If a child turns 8 while in the study – will they keep using the cargiver as proxy or switch to self-report? 11. One of the implementation outcomes for success is that phone reminders are not required. Please provide justification for this choice. Elsewhere it seems that families had a choice to complete the PROMs in clinic before their visit - I would imagine that most families would require phone reminders (or just ignore them and complete when they get to the office). It also could be that families from lower SES environments may have less access to technology and are unable to complete at home. 12. It was not clear if the study team was including clinicians in the interviews to determine if they also felt the intervention was successful.

REVIEWER	Ashima Singh Medical College of Wisconsin, USA
REVIEW RETURNED	03-Aug-2020

GENERAL COMMENTS	The authors present the protocol of an interesting study aimed to evaluate the use of Patient reported outcomes among children with skin conditions in clinic settings. Below are some suggestions to further clarify the manuscript. 1. Introduction: end of first paragraph. the comment about the study including parent proxy data needs a better connector or else is confusing the way it is now. Child self-report and parent-proxy collection of data should be clearly mentioned in the methods, and maybe is not needed within the introductory paragraph. 2. Aims and objectives: Suggest deleting the headers of effectiveness and implementation ‘outcomes’. For the effectiveness secondary aim: what are the other health-related quality outcomes of interest. Please specify.

	3. In the hypotheses for the effectiveness aim: please specify the comparator. How is point 1 different from point 2 – please clarify. 4. Specify hypothesis for the implementation aim. 5. Page 10, lines 8-10: “Theory-based interventions also tend to be more effective than non-theory based interventions” reads out of context. Maybe more suited in discussion section. 6. Specify the age of children recruited in methods. 7. Consolidated framework is a well accepted method for conducting implementation research. However, it is unclear how the outcomes are mapped to the framework – it will be helpful to explain how the information for specific CFIR domains is collected and how those link with the implementation outcomes. Also, the implementation evaluation acceptability is assessed by objective compliance of appointments. Was CFIR framework used only to assess sustainability? More explanation will be helpful. 8. Sample size determination – Is 0.68 a meaningful change in quality of life. 9. Is fidelity considered an implementation outcome or an effectiveness outcome? Please put it in context. 10. Please elaborate the reasons for not administrating the exact same set of instruments for the two sites participating in the study.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1
Comment: There were a number of places where the protocol write-up could be streamlined and some of the choices better justified.
Response: We have streamlined the protocol in several areas of the manuscript and added further justification regarding our choices, which we hope has improved the manuscript.

Comment:
I have include specific questions below:
1. Hypotheses 2: The authors mention the PEDS-ePROM and ePROM but both of these terms were not fully defined until later in the paper.
Response: Thank you for picking up this oversight. We have now defined these terms early in the manuscript.

Comment:
2. It is not clear if the authors considered the impact of combining proxy and self-report into the same sample. It has been shown that agreement between children and caregivers is often low on PROMs. More justification for this choice should be provided.
Response: We did not consider combining the proxy and self-report as agreement between children and caregivers is often low as highlighted by reviewer 1. We have added further justification for our choices. The primary outcome will be based upon proxy caregiver-report which we have made clearer throughout the manuscript.

Comment:
3. It was not clear if randomization would be done within each target population (burn scars vs haemangiomas). What do the authors expect to be the final breakdown of the sample? (ex. 50% burns and 50% haemangiomas?) This has implications for the secondary data analysis.
Response: Randomisation has been conducted within each target group which we have clarified in the manuscript. At the commencement of the randomised controlled trial we expected the breakdown of the samples to be approximately equal. However the impact of COVID-19 on health service activity

could result in an unanticipated imbalance in the target populations.

Comment:
4. Do the authors expect severity of HRQoL to be different among the two groups of patients?
Response: These two groups have not been compared in terms of health-related quality of life to our knowledge. However based on our clinical knowledge we expect similar impacts on generic health-related quality of life as proxy-reported impairments in psychosocial functioning and caregiver distress are common to both conditions, due to appearance concerns and stigma from people in the community. If we assume that both groups are receiving the most appropriate intervention at the time the PEDS-ePROM intervention commences, then we can reasonably expect that improvements in health-related quality of life would be similar. The physical conditions also have a similar trajectory of symptom improvement over a 1-year period with usual care intervention. We will report the impact of the conditions on health-related quality of life for the two conditions at baseline.

Comment:
5. Would the authors expect the intervention to be better received by one group or the other?
Response: Thankyou for highlighting this important point. We do not expect the intervention to be better received by one group than the other based on the interviews we conducted in the design stages of the project, although it is possible this might be the case. Feedback of patient-reported outcome measure results using graphical displays was not occurring routinely for either group at the commencement of the study. Perceptions regarding how well the intervention is received by children and caregivers, as well as by clinicians will be confirmed using interviews and reported as part of the implementation component. Interviews in the pre-implementation stage of the study indicated the intervention was generally viewed positively by children and caregivers in both groups but not by all clinicians.

Comment:
6. It was not clear why “referrals” were chosen as an outcome. Is a higher number of referrals a positive or a negative?
Response: Referrals were chosen as an outcome as it was expected that clinicians would ask children and caregivers about results provided in the graphical displays during routine consultations, and if there were psychosocial concerns a referral would be made to another health professional such as a social worker who did not typically attend the outpatient clinics. A higher number of psychosocial referrals was expected in the group receiving the graphical display of results, as psychosocial concerns were expected to be identified and raised in consultations more often than the group who completed PROMs but had no results sent into consultations. Our stance is consistent with a recent systematic review by Bele (2020) involving caregivers of children with chronic conditions whereby the authors also expected a higher number of referrals in the intervention group compared to the control groups to be a positive outcome. In that review three of the six included studies evaluated the impact of PROMs on the referral rate. Two of the studies reported an increase in the referral rates in the intervention group and one study identified no difference in referral rates between intervention and control groups. However, contrary to the authors of the review, the two caregiver consumers involved in the systematic review perceived a positive outcome to be reduced referrals to health professionals as a result of receiving PROM feedback interventions. As mixed findings have been identified from previous research, we will seek to clarify understanding of this outcome in interviews. We have added further justification for this outcome.

Comment:
7. It was not clear why treatment satisfaction was chosen as an outcome.
Response: Treatment satisfaction was included as an outcome as it was expected that children and caregivers who had graphical results and priorities raised in consultations would be more satisfied with their treatment. This was based on the finding that significantly more intervention patients

reported satisfaction with overall care in a study of children with diabetes, which was the only paediatric study that examined this outcome in the systematic review by Ishaque et al. 2019. A rationale for the inclusion of this outcome has been added to the manuscript.

Comment:
8a. Are the authors planning to stratify randomization among age groups? We know that age makes a difference in HRQoL reporting and again there are implications for combining proxy vs self-report. It would strengthen the paper if the authors could provide justification around their choice.
Response: This is another important point. Randomisation was not stratified by age group due as the size of some of the groups would have been relatively small. We will consider this for a larger trial that is planned. For the current study we will examine baseline differences in health-related quality of life reported by age (e.g., 0-12 months; 1-2 years; 2-8 years; 8+). If baseline differences in overall health-related quality of life (proxy-reported) are present, age will be controlled for in the linear mixed models analysis.
Child-reported health-reported quality of life data will be examined in secondary analyses so child and proxy report data will not be combined. Proxy-report will be used as the primary outcome as parents' perceptions of children's quality of life influence outcomes such as healthcare utilisation (Rodday et al. 2017), and later child outcomes (Koot & Verhulst 1992; Verhulst et al. 1994).

Comment:
8b. Inclusion criteria – a lower bound for age is not specified. Are the authors planning to recruit infants?
Response: A lower bound for the age of participants has been added. Infants have been recruited.

Comment:
9. Why does the study team believe that the graphical display will significantly improve outcomes over just receiving the intervention?
Response: The graphical displays are expected to significantly improve outcomes over just completing PROMs with no feedback of results. This expectation is based on the systematic review by Ishaque et al. 2019 that identified PROM interventions worked better when PROM results were provided to clinicians compared to when they were not provided to clinicians based on significant differences from robust and non-robust studies. In the busy clinic environments, parents raising concerns that may be identified using PROMs completion alone can be challenging as multiple medical or allied health professionals are often present reviewing the child, and the environment is not conducive to raising psychosocial concerns in particular. We have added further background information to make this clearer.

Comment:
10. More information is needed around the power calculation and sample size justification.
a. How was the effect size chosen? What metric is the effect size reported in? Is this for the primary outcome (PedsQL)?
b. No attrition or loss of follow-up was taken into account. What happens if a number of people drop out?
Response:
We have added additional information regarding the power calculation and sample size justification. The effect size was based on the number of participants who could feasibly be recruited and randomised in the 6-month intervention period with 80% power, alpha 0.05. A between group difference of 0.68 standard deviations is considered clinically meaningful at the individual level by expert clinicians.

Our sample size estimate was based on outcome data being available for 70 participants. We have added additional information on attrition based on prior studies in the study setting to indicate that of

those randomised we expect approximately 10 percent attrition at 3-month follow-up and approximately 20 percent attrition at 6-month follow-up. We will continue recruitment until we have recruited 88 participants to account for expected attrition of 20 percent at 6-month follow-up.

Comment:
11. Will the study team plan to take into account intervention 'dosage'? (E.g. potential differences if child gets PROM feedback 6 times over 6 months versus 2 times).
Response: We will conduct a secondary analysis to calculate the maximum potential effect of the intervention with children analysed according to the treatment actually received (that is, we will conduct an 'as treated' analysis incorporating treatment dose received). This information has been added to the effectiveness analysis section of the manuscript.

Comment:
12. If a child turns 8 while in the study – will they keep using the caregiver as proxy or switch to self-report?
Response: If a child turns 8 while in the study they will keep using the caregiver as proxy. We have added this detail to the inclusion criteria.

Comment:
13. One of the implementation outcomes for success is that phone reminders are not required. Please provide justification for this choice. Elsewhere it seems that families had a choice to complete the PROMs in clinic before their visit - I would imagine that most families would require phone reminders (or just ignore them and complete when they get to the office). It also could be that families from lower SES environments may have less access to technology and are unable to complete at home.
Response: We have added a justification for the choice of phone reminders not being required as a successful outcome. This choice was based on there being a lower burden on clinic staff and better uptake of the intervention in routine clinical practice if at least a proportion of the patient-reported outcome measures are completed without the need for reminder phone calls.

Comment:
14. It was not clear if the study team was including clinicians in the interviews to determine if they also felt the intervention was successful.
Response: Clinicians are being included in the interviews to determine if they also feel the intervention implementation is successful and well targeted from a clinician perspective. This details has been made clearer in the manuscript.

Reviewer: 2

Comment:
1. Introduction: end of first paragraph. the comment about the study including parent proxy data needs a better connector or else is confusing the way it is now. Child self-report and parent-proxy collection of data should be clearly mentioned in the methods, and maybe is not needed within the introductory paragraph.
Response: Thank you for this suggestion. The comment about including parent-proxy data has now been removed from the introduction but has been clearly mentioned in the methods section.

Comment:
2. Aims and objectives: Suggest deleting the headers of effectiveness and implementation 'outcomes'. For the effectiveness secondary aim: what are the other health-related quality outcomes of interest. Please specify.
Response: The headers of effectiveness and implementation have been removed from the aims and

objectives section. For the effectiveness secondary aim we have now specified the other health-related quality of life outcomes.

Comment:
3. In the hypotheses for the effectiveness aim: please specify the comparator. How is point 1 different from point 2 – please clarify.
Response: We agree with the reviewer that the difference between the hypotheses as they were written was not clear thus the hypotheses 1 and 2 have been combined into a single hypothesis. We have added an additional hypothesis regarding referrals and treatment satisfaction to make our expected outcomes clearer.

Comment:
4. Specify hypothesis for the implementation aim.
Response: We have not specified implementation hypotheses as we do not know what we expect to find. Not including implementation hypotheses is consistent with implementation process evaluation methodology which differs to a formative evaluation where a priori aims and hypotheses are made and formal feedback of the implementation findings provided during the study (Bauer et al, 2015). Further, our implementation findings are likely to be complex (multilayed and multifaceted) based on our theoretically-driven study, which would be difficult to capture in a single hypothesis or series of hypotheses.

Comment:
5. Page 10, lines 8-10: “Theory-based interventions also tend to be more effective than non-theory based interventions” reads out of context. Maybe more suited in discussion section.
Response: As suggested this point has been moved to the discussion.

Comment:
6. Specify the age of children recruited in methods.
Response: The age of the children recruited has been added in the methods.

Comment:
7. Consolidated framework is a well accepted method for conducting implementation research. However, it is unclear how the outcomes are mapped to the framework – it will be helpful to explain how the information for specific CFIR domains is collected and how those link with the implementation outcomes. Also, the implementation evaluation acceptability is assessed by objective compliance of appointments. Was CFIR framework used only to assess sustainability? More explanation will be helpful.
Response: Thank you for highlighting the need to further explain this method. Additional information has been added to the introduction and the methods sections to make the link between the Consolidated Framework for Implementation Research and implementation outcomes clearer.

Comment:
8. Sample size determination – Is 0.68 a meaningful change in quality of life.
Response: An effect size of 0.68 standard deviation units is the between-group mean difference in overall health-related quality of life at 6 month follow-up rather than the change in health-related quality of life. Clinical experts we have consulted believe that a difference of 0.68 standard deviations is a clinically important difference.

Comment:
9. Is fidelity considered an implementation outcome or an effectiveness outcome? Please put it in context.
Response: Fidelity is considered an implementation outcome. This has been made clearer and

additional detail added to Table 2.

Comment:
10. Please elaborate the reasons for not administrating the exact same set of instruments for the two sites participating in the study.
Response: The exact same set of instruments were not administered to the two sites as disease-specific measures have been documented as being more sensitive to change than generic measures of health-related quality of life. Thus condition-specific measures are being completed in each clinic in addition to the same generic measures across both clinics. This rationale has been added to the manuscript.

References:
Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. BMC Psychol. 2015;3(1):32.
Verhulst FC, Koot HM, Van der Ende J. Differential predictive value of parents' and teachers' reports of children's problem behaviors: A longitudinal study. J Abnorm Child Psychol. 1994; 22:531–546.
Koot, H. M., & Verhulst, F. C. (1992). Prediction of children's referral to mental health and special education services from earlier adjustment. Child Psychology & Psychiatry & Allied Disciplines. 33(4), 717–729.
Rodday AM, Graham RJ, Weidner RA, et al. Predicting Health Care Utilization for Children With Respiratory Insufficiency Using Parent-Proxy Ratings of Children's Health-Related Quality of Life. Journal of Pediatric Health Care. 2017; 31(6):654-662.
Ishaque S, Karnon J, Chen G, et al. A systematic review of randomised controlled trials evaluating the use of patient-reported outcome measures (PROMs). Qual Life Res 2019;28(3):567-92.

VERSION 2 – REVIEW

REVIEWER	Christina Zigler Duke University School of Medicine
REVIEW RETURNED	11-Nov-2020

GENERAL COMMENTS	The study aim is to investigate: (1) the effectiveness of a patient-centred care intervention using feedback from electronic patient-reported outcome measures (PROMs; PEDS-ePROM intervention) on health outcomes, referrals, and treatment satisfaction; and (2) the implementation of PEDS-ePROM by assessing acceptability and sustainability of the intervention and study processes. The first aim looks to see if the intervention improved health-related quality of life (HRQoL; primary outcome), scores on other PROMs, referrals, and overall tx satisfaction. The second aim looks at the overall implementation process of the intervention. From the perspective of the child and caregiver, they will explore acceptability and sustainability. There are a number of strengths of the study. They utilize a pragmatic randomized-controlled trial (varying presence of the PROM graphical display) and mixed methods with quantitative and qualitative data being collected from multiple sources (child, caregiver, EHR, clinicians). They also include a battery of PROMs, including generic scales and disease-specific measures. The authors were responsive to a number of critiques after initial review, however, a few concerns remain. There was no explicit response to reviewer comments, which would have helped identify justification around their responses. Major concerns:

	The authors seem to state that only proxy measures will be used bc 'young children' cannot self report. This is not true, as young children can self-report using measures designed for them. Additionally, PedsQL does offer child direct report after age 5. In contrast to the initial wording, on page 14, the protocol states that children 0 to 16 years will be enrolled. Thus, the protocol still does not address how caregiver proxy and child self-report will be integrated and their justification for doing so. We know from the literature that often parents and children do not agree on their reports of quality of life, so combining them into one analysis will require convincing justification. No clear if there will be any sampling based on age to ensure representation across the life span. I'm still unsure what PEDS-ePROM and ePROM mean based on page 6. Is PEDS-ePROM the intervention study itself? And ePROM is the graphical display of information? Later on page 7, it becomes clear that PEDS-ePROM is the condition of PROMs + graphical display, where ePROM is just the PROMs being administered without graphical display. So they are intervention conditions (group 1 vs group 2). On page 13, the intervention groups are described again but not referred to using acronyms. On page 15, state they exclude non-English speakers because "PROMS are only available for the study in English". But the PedsQL is available in multiple languages, so this is not sufficient justification. https://www.pedsql.org/translations.html. Some questions remain around the sample breakdown: Do the authors expect severity of HRQoL to be different among the two groups of patients? Would the authors expect the intervention to be better received by one group or the other?

REVIEWER	Ashima Singh Medical College of Wisconsin
REVIEW RETURNED	02-Nov-2020

GENERAL COMMENTS	Thank you for addressing the comments.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Comments to the Author
The study aim is to investigate: (1) the effectiveness of a patient-centred care intervention using feedback from electronic patient-reported outcome measures (PROMs; PEDS-ePROM intervention) on health outcomes, referrals, and treatment satisfaction; and (2) the implementation of PEDS-ePROM by assessing acceptability and sustainability of the intervention and study processes. The first aim looks to see if the intervention improved health-related quality of life (HRQoL; primary outcome), scores on other PROMs, referrals, and overall tx satisfaction. The second aim looks at the overall implementation process of the intervention. From the perspective of the child and caregiver, they will explore acceptability and sustainability.

There are a number of strengths of the study. They utilize a pragmatic randomized-controlled trial (varying presence of the PROM graphical display) and mixed methods with quantitative and qualitative data being collected from multiple sources (child, caregiver, EHR, clinicians). They also include a battery of PROMs, including generic scales and disease-specific measures.

1. **COMMENT:**
The authors were responsive to a number of critiques after initial review, however, a few concerns remain. There was no explicit response to reviewer comments, which would have helped identify justification around their responses.

RESPONSE: Thank you for pointing out the strengths of the study. We have attempted to rectify the lack of sufficient detail in our previous responses to some of the reviewer comments in areas of concern.

2. **COMMENT:**
Major concerns:
The authors seem to state that only proxy measures will be used bc 'young children' cannot self report. This is not true, as young children can self-report using measures designed for them. Additionally, PedsQL does offer child direct report after age 5.

RESPONSE: We have amended any statements that suggest that young children age 5 years and older cannot self-report which we did not mean to imply. We meant that very young children are unlikely to be able to self-report based on their young age or cognitive capacity and as some of the chosen measures were not developed or validated using the self-report of very young children. We have chosen the cutoff of 8 years for proxy versus child report for all measures based on our development and validation of burn-scar specific measures and research experience using generic measures with our burn scar population that has indicated children younger than 8 years often struggle to comprehend the included measures of health-related quality of life [1]. Our age cut-off aligns with that of Arbuckle et al (2013) who identified the strongest support was for the broad age-range of 6-8 years being the youngest age children can meaningfully report on a patient-reported outcome [2]. It also aligns with the proxy versus child report age-cut-off for the administration of health-related quality of life measures in routine clinical practice in the burn scar clinic setting in our study. For our infantile haemangioma population children are expected to be infants and toddlers, thus cannot self-report using ePROMs thus only caregiver proxy report will be used.

We have added further justification for our 8-years of age cut-off of proxy-report versus child report outcomes to the manuscript. This includes support for the comparability of parent-proxy and child self-report in children who have sustained burns. A recent systematic review of health-related quality of life in children with burns identified that parent-proxy and child self-ratings were generally comparable based on generic and burn specific measures [3]. This finding is supported by an additional two trials examining burn scar specific health-related quality of life that have been conducted in our burn scar clinic [4,5]. These trials identified health-related quality of life as similar across parent proxy and child report. Thus the analysis population will consist of all participants who have analysable data. To investigate possible effects of informant, a pre-specified subgroup analyses of the primary and secondary outcomes of health-related quality of life will be stratified by informant (proxy versus child report), in addition to age and gender to determine whether effect differences exist based on these factors. A sensitivity analysis will be conducted to compare the results of the parent proxy versus child self-report where available. The stratified analyses will assist in determining whether the informant has any effect on specific subscales in the secondary outcome measures such as appearance and family disruption [3].

3. COMMENT:
In contrast to the initial wording, on page 14, the protocol states that children 0 to 16 years will be enrolled. Thus, the protocol still does not address how caregiver proxy and child self-report will be integrated and their justification for doing so. We know from the literature that often parents and

children do not agree on their reports of quality of life, so combining them into one analysis will require convincing justification.

RESPONSE: Please refer to our response to comment 2 regarding the parent proxy and child self-report.

COMMENT:
No clear if there will be any sampling based on age to ensure representation across the life span.

RESPONSE:
There will be no representative sampling across the lifespan. However the sample in our randomised controlled trial (RCT) should reflect the clinical population as we will consecutively sample eligible participants. A previous study by the author team using the same consecutive sampling in the study setting with the same population [4] demonstrated representation of the study population [6]. In addition, we will report the results overall as well as stratified for age for the RCT.

For the implementation component of the study we will use purposeful sampling to obtain in-depth information on the implementation in both intervention arms. Further details have been added to the methods to detail these aspects.

4. **COMMENT:**
I'm still unsure what PEDS-ePROM and ePROM mean based on page 6. Is PEDS-ePROM the intervention study itself? And ePROM is the graphical display of information?

RESPONSE:
We apologise for the confusion. We have now removed the intervention names (but kept a broad study name) and hope this has made the reporting clearer.

5. **COMMENT:**
Later on page 7, it becomes clear that PEDS-ePROM is the condition of PROMs + graphical display, where ePROM is just the PROMs being administered without graphical display. So they are intervention conditions (group 1 vs group 2).

RESPONSE: We have removed the abbreviations for the intervention groups as per comment 5.

6. **COMMENT:**
On page 13, the intervention groups are described again but not referred to using acronyms.

RESPONSE: We have addressed the inconsistency on page 13 by removing the use of acronyms to refer to intervention groups.

7. COMMENT:
On page 15, state they exclude non-English speakers because "PROMS are only available for the study in English". But the PedsQL is available in multiple languages, so this is not sufficient justification. https://www.pedsql.org/translations.html.

RESPONSE:
We agree that some of our chosen PROMs are available in languages other than English. The PEDS-QL, CHU-9D and Brisbane Burn Scar Impact Profile were available in some of the languages spoken by potential non-English speaking participants. However as highlighted by Ravens-Sieberer et al. (2006) even though measures may be available in different language few have been developed in a culturally appropriate manner and psychometrically tested for those groups [7]. With regard to the PEDS-QL, which we used as our primary outcome, Stevanovic et

al. (2015) found cross-cultural measurement variance for the PedsQL adolescent self-report thus it was recommended that caution be used in making cross-cultural quality of life comparisons using that measure [8]. In addition, disease-specific PROMs were either not available in other languages or only in a small number of languages.

We had other pragmatic reasons for not including people who could not speak English. This included our study operating on a limited budget without the capacity to fund the licences for multiple translations of the PEDS-QL (US$500 per additional translation in addition to English) and being unable to anticipate which translations and how many translations would be required prior to the commencement of the study. We also had no specific funding for the transcription of interviews that might require an interpreter and would need non-English transcription services. The other reason we did not include these participants was that the intervention was not developed with non-English speaking participants. Thus we were unable to ensure the intervention would meet the needs of this vulnerable group consistent with the World Medical Association Declaration of Helinski: "*Medical research with a vulnerable group is only justified if the research is responsive to the health needs or priorities of this group and the research cannot be carried out in a non-vulnerable group. In addition, this group should stand to benefit from the knowledge, practices or interventions that result from the research" (p. 1291)*[9]. We will ensure that diverse cultural groups will stand to benefit from the study through the implementation component of the study which will explore how the needs of people with difficulty speaking or understanding written English can best be addressed in the future.

We have updated our rationale for excluding participants who were unable to speak English which we have acknowledged as a limitation and have clearly indicated that our implementation component of the study will address this area. We will report the number of participants excluded for this reason in the study reporting.

8. **COMMENT:**
Some questions remain around the sample breakdown: Do the authors expect severity of HRQoL to be different among the two groups of patients? Would the authors expect the intervention to be better received by one group or the other?

RESPONSE: We have tried to improve our response to both of these points providing further justification for what we expect and what we will do if unexpected differences emerge for these factors.

We expect a similar mean severity of health-related quality of life in the two groups of patients who are usually treated for a similar length of time, receive more intensive intervention initially if indicated, and in whom appearance-related concerns by parents can be heightened. Both groups are likely to include a range of severities of health-related quality of life based on our prior research with these groups. Both groups are similar in that skin conditions affecting the vascular system are the primary focus of treatment in consultations in the study setting.

To account for the possibility of imbalance in mean severity of health-related quality of life in the two clinics for the RCT, we will conduct a sensitivity analysis reporting the results overall across the clinics as well as after stratification for clinic for generic health-related quality of life and will report any imbalance in mean severity of health-related quality of life at baseline. We have updated the methods section of the protocol to outline these analyses.

Our pre-implementation interviews identified that the intervention appeared to be well received by both groups of patients thus we did not expect the intervention to be better received by one group or the other. However, our implementation evaluation will determine how well the intervention is received by children and caregivers, as well as by clinicians in the implementation phase using

interviews. Rating coded data using a positive to negative valence scale as part of acceptability will be part of this evaluation. Using purposeful sampling we will interview families in whom the intervention was well received and not well received to gain further insights regarding how well the intervention was received. We have updated parts of the methods section to make these details clearer.

References

1. Tyack Z, Ziviani Z, Kimble R, Plaza R, Jones A, Cuttle L, Simons M. (2015). Measuring the impact of burn scarring on health-related quality of life: development and preliminary content validation of the Brisbane Burn Scar Impact Profile (BBSIP) for children and adults. *Burns, 41 (7), 1405 - 19*.
2. Arbuckle, R., Abetz-Webb, L. "Not Just Little Adults": Qualitative Methods to Support the Development of Pediatric Patient-Reported Outcomes. *Patient* **6,** 143–159 (2013). https://doi.org/10.1007/s40271-013-0022-3
3. Spronk I, Legemate CM, Polinder S, van Baar ME. Health-related quality of life in children after burn injuries: A systematic review. Journal of Trauma and Acute Care Surgery. 2018;85(6).
4. Wiseman J, Ware RS, Simons M, et al. (2020). Effectiveness of topical silicone gel and pressure garment therapy for burn scar prevention and management in children: a randomized controlled trial. *Clinical Rehabilitation*. 2020;34(1):120-131. doi:10.1177/0269215519877516
5. Frear CC , Cuttle L, McPhail SM, Chatfield MD, Kimble RM, Griffin BR. Randomized clinical trial of negative pressure wound therapy as an adjunctive treatment for small-area thermal burns in children. Br J Surg. 2020 Dec; 107(13): 1741–1750.
6. Stockton KA, Harvey J, Kimble RM. (2015). A prospective observational study investigating all children presenting to a specialty paediatric burns centre. *Burns.* 2015; 41(3), 476-483.
7. Ravens-Sieberer U, Erhart M, Wille N, Wetzel R, Nickel J, Bullinger M. Generic health-related quality-of-life assessment in children and adolescents: methodological considerations. Pharmacoeconomics. 2006;24(12):1199-220. doi: 10.2165/00019053-200624120-00005
8. Stevanovic, D., Atilola, O., Vostanis, P., Pal Singh Balhara, Y., Avicenna, M., Kandemir, H., Knez, R., Franic, T., Petrov, P., Maroco, J., Terzic Supic, Z. and Bagheri, Z. (2016), Cross-Cultural Measurement Invariance of Adolescent Self-Report on the Pediatric Quality of Life Inventory™ 4.0. J Res Adolesc, 26: 687-695.
9. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191-2194.

VERSION 3 – REVIEW

REVIEWER	Christina Zigler Duke University School of Medicine
REVIEW RETURNED	27-Jan-2021

GENERAL COMMENTS	The authors were extremely responsive to reviewer comments and their justification for choices are well described in the updated manuscript.