Author’s response to reviews

Title: Initial Development of Supportive Care Assessment, Prioritization and Recommendations for Kids (SPARK), a symptom screening and management application

Authors:

Sadie Cook (sadie.cook@sickkids.ca)
Emily Vettese (emily.vettes@sickkids.ca)
Dilip Soman (dilip.soman@rotman.utoronto.ca)
Shannon Hyslop (shannon.hyslop@sickkids.ca)
Susan Kuczynski (liaison@opacc.org)
Brenda Spiegler (brenda.spiegler@sickkids.ca)
Hailey Davis (haileydavis@live.com)
Nathan Duong (nathan.225@hotmail.com)
Stacee Ou Wai (stacee@translucentcomputing.com)
Robert Golabek (robert@translucentcomputing.com)
Patryk Golabek (patryk@translucentcomputing.com)
Adam Antoszek-Rallo (info@catalystworkshop.com)
Tal Schechter (tal.schechter-finkelstein@sickkids.ca)
Lee Dupuis (lee.dupuis@sickkids.ca)
Lillian Sung (lillian.sung@sickkids.ca)

Version: 1  Date: 09 Oct 2018

Author’s response to reviews:

Lillian Sung MD, PhD
Division of Haematology/Oncology
The Hospital for Sick Children
October 9, 2018

Dear Dr. Krüger,

We wish to thank you and the reviewers for the careful review of our manuscript entitled “Initial Development of Supportive Care Assessment, Prioritization and Recommendations for Kids (SPARK), a symptom screening and management application”.

The comments were very helpful and we believe that the manuscript has been greatly improved by the modifications. Please find below an itemized list of the comments and how they have been addressed in the manuscript. We hope that you will find these modifications satisfactory.

Thank you for your time and consideration. We look forward to further correspondence with your office.

Sincerely,

Lillian Sung, MD, PhD
Reviewer 1:

1. The paper is important and clearly demonstrates the iterative nature of health app co-design that many health technologies should employ. The paper can be improved by clearer depictions of the three different testing phases through diagrams and subheadings to take the reader along the development journey.

Response: Thank you for your review. We have incorporated your feedback as described below.

2. A flowchart (perhaps of screenshots) of the methods described in lines 124-133 would provide more clarity. This would enable the reader to understand the flow of the app more clearly which is critical to understanding this paper.

Response: A flow diagram has been added as Figure 3 to illustrate the flow of the participant interview.

3. The thinkaloud process is referred to as an interview, however it doesn't seem as if it was an interview. Were participants asked specific questions? Inclusion of an Interview guide may be useful.

Response: The thinkaloud process was incorporated within an interview. To clarify how interviews were conducted, we added the following to the Methods:

“An interview guide was used to ensure consistency in the approach.”

“The interview guide began by asking the participant to explore a specific section of the website and to explain what they were seeing and their reactions during this navigation process. This general approach was then followed by specific questions asking the participant to complete specific tasks or their understanding of specific elements.”

4. Why were the materials in the 'low-fidelity phase' grey scale?

Response: Material in the low-fidelity phase were grey scale because the purpose was to evaluate function and whether participants understood how to navigate the website. Color adds another dimension of preferences and thus, this addition was delayed until the design phase. In order to address this comment, we have added the following to the Methods:

“All pages were in black and white (grey scale) to determine whether participants understood how to use SPARK without distraction by color, since addition of color would introduce preferences for a specific design.”
5. For the Procedures section, it might be easier to use subheadings to categories what was assessed e.g. Visual preferences of: icon, etc. Understanding of: xyz. It is a bit unclear presently.

Response: We appreciate this comment. We presented the Procedures section by phase of development and since almost all of the elements were evaluated in both low fidelity and high fidelity, we could not add subheadings to this section. However, to try and make this clearer, we added the following to the Methods:

“To make it clear what sections were specifically evaluated for understanding, the “label” used in the Results follows the task description in bold and parenthesis.”

An example of this can be found as follows (which we believe makes the linking between the Methods and Results more clear:

“After free-exploring, we asked participants how they would perform specific tasks such as completing SSPedi (Completing SSPedi) or seeing their current or previous SSPedi scores (Seeing current or previous SSPedi score).”

6. The Results section will be a lot easier to follow with subheadings, and perhaps a table if suitable (for the results that involve %’s).

Response: Subheading were added to the Results section.

7. The Discussion does not really discuss results. Further interpretation of the findings of user preferences and how these can be used as lessons for other health app developers would be useful.

Response: Thank you for your feedback. To address this comment, we added the following to the Discussion:

“We found this iterative and phased approach allowed participants to focus on individual aspects (for example navigation or design), and provided valuable feedback on distinct elements. This approach to measuring user preferences may be useful for other health app developers.”
Reviewer #2:

8. This manuscript represents a very relevant study that required a lot of investment in terms of time, energy and resources by the research team for both planning and realizing the project. The methodology seems to be well thought-out and combines several steps and an iterative work. The participants sample is impressive given the context of the care delivery chosen, the nature of the disease and the tasks required, enrollment of 90 children and two interviewers to conduct all assessments, Invision platform use... The team research used a co-construction approach involving researchers, patients, and technology developers that contributed to a relevant product, most likely useful and usable for sick kids, their families and health professionals.

Response: Thank you for your review and comments

9. There are a lot of acronyms used throughout.

Response: We reduced the use of acronyms throughout and only retained two key acronyms which relate to the website (SSPedi and SPARK).

10. To present the proposed items and give more information about measurement tools development and validation (measure of Understandability, Usability and Ease of understanding of SPARK reports concepts)?

Response: The use of Likert-type scales is commonly used across instrument development and evaluation of patient-reported outcomes. We have previously used these measures of understandability, usability, ease of use and usefulness in prior studies (1-3) and although we have not formally evaluated their psychometric properties, they have face validity. We have found them easy to implement and interpret.

11. Explain why the authors choose that the Understandability of SPARK components were evaluated concept it has 4-point Likert scale while the two other concepts (Usability and Ease of understanding of SPARK reports was evaluated on a 5-point Likert scale?

Response: Lines These instruments were developed independently and have been used in several studies as described in Comment #10. Given our extensive experience with both scales, we did not wish to modify one of them at this point and modification would not have served a functional purpose.
12. In the discussion section, authors talk about "Utility of SPARK from the perspective of children with cancer", this aspect was not evaluated in this study?

Response: We agree that this was confusing. When describing utility, we are referring to children’s perception of the usefulness of SPARK for future children receiving cancer treatments. To make this clearer, we have replaced the word “utility” with “usefulness” in the Discussion.

13. Also, it would be appropriate to specify in the conclusion that the platform has been evaluated with children 8 years and older for the studied aspects.

Response: We agree and have made this modification to the Conclusion as follows:

“In conclusion, SPARK is a web-based application which is usable and understandable from the perspective of children aged 8 to 8, receiving cancer therapies.”

14. Comment: On page 10, missing the word "Help" in the sentence - "understood the "How Will SSPedi Me?", bottom of page.

Response: Thank you – this has been corrected.

Reviewer #3:

15. Strengths:

The paper is well written.

Problems, motivations, and objective that justified a need of study are clearly presented in the Background section.

The development and evaluation phases have been performed in well-structured procedures, and the results are interesting and valid.

Response: Thank you for your feedback.

16. The paper is focused on presenting proposed methods but failed in showing the earlier work done, related concepts, and contexts. Please review some current works done on
development and evaluation of HIS/web-based clinical system in the Materials section. There are many works in this area, and you should compare their methods with your proposed one and justify why your method is more significant and valid.

Response: We agree with the need to show more information in the Methods. To address this comment, the following was added:

“Theoretical Background: We used the Technology Acceptance Model as the guiding principle behind this research. This model states that perceived ease of use and perceived usefulness are predictors of behavioral intention and actual use of health technology(4). Consequently, we focused on evaluating these domains during SPARK development.”

In terms of justifying why our method is more significant and valid, we don’t believe this is true. Rather, we relied upon existing methods and frameworks to evaluate our product in a novel circumstance, namely children as the user-patients.

17. The novelty of the proposed method is somewhat unclear.

Response: There are three aspects of this manuscript which are novel. First, the product itself is innovative. There are no products to facilitate symptom screening and access to guidelines for children receiving cancer treatments. Second, use of children with a serious medical condition as the user-patient is novel. Third is the development of a Review Panel that includes patient representatives. We have addressed this issue with the addition in the Discussion:

“We successfully showed how we could use pre-existing methods and using the Technology Acceptance Model, engage with a pediatric population with a serious medical condition to refine a health technology. Since SPARK is ultimately intended toward this population, it was important that understandability, usability and ease of use be evaluated from their perspective.”

“A novel aspect of this study was the use of a Review Panel with patient representation and inclusion of a behavioral scientist specifically focused on choice architecture to enhance SPARK design and usability.”

18. Why have you not standardized the use of Likert scale like a 4 point scale for quantitative evaluation of SPARK?

Response: Please see response to Comment #11.
19. What is the basis to select these questions for quantitative evaluation of SPARK? Are the questions adopted from TAM? Are the questions chosen related to technical problems that required for SPARK design.

Response: Please see response to #16.

20. You can improve the reliability and validity of quantitative evaluation results. You must provide some more details about the reliability and validity results for every tested question.

Response: Please see response to Comment #11. In terms of patients’ assessment of ease of use and usefulness, the use of a simple Likert scale in this circumstance has face validity. The one item that we were concerned about in terms of measurement was the interviewers’ assessment of understandability and thus, we used two evaluators for this item. To improve our explanation of how this was accomplished, we added the following to the Methods:

“The two evaluators rated understandability independently and then compared ratings. If they disagreed, they referred to field notes to arrive at consensus.”

21. How do you interpret the overall results? Is it highly accepted after SPARK improvement? Do you conduct follow-up or post-test after addressing all participants’ feedbacks?

Response: In the methodology we used, we continued to recruit new patients until SPARK was considered acceptable based upon a priori established criteria. However, the Reviewer raises an excellent point, namely that further evaluations may be useful. In order to address this comment, we added the following to the Discussion:

“Finally, further evaluation of the refined version of SPARK would be useful and we intend to include further qualitative evaluations during longitudinal evaluation to identify whether further modifications would be beneficial.”

22. The section is more on supplying the benefits from design and evaluation of SPARK. You should justify why your result is significant and highly accepted by comparing to the earlier works.

Response: Please see responses to Comments #17 and #21.

23. Have your study offers any commercial or societal impact? I am not clear of your study contributions to both theory and clinical practice, but the practical context is somewhat
missing. To address the study limitations for better generalizability of findings, your conclusion also should provide solutions for future researchers in engaging more participants from different clinics by utilizing a random sampling technique.

Response: Please see response to Comment #17 regarding importance. We agree that including patients from different centers would improve generalizability and thus, have added the following to the Discussion:

“Enrollment of participants from different centers would have improved generalizability but would have made the iterative evaluative process more logistically complex.”

Reviewer #4:

24. Supportive care Prioritization, Assessment and Recommendations for Kids (SPARK) is a web-based application about the symptom screening and the supportive care for pediatric patients. This paper presents interesting and original article but some comments about the paper that I hope will help to improve it.

Response: Thank you for your feedback.

25. Please describe other limitations besides linguistic limitations (only with English-speaking children), which is described in the discussion section. Please define the limitations of various aspects (development and user aspects), more clearly.

Response: We have added the following to the Discussion to address this comment:

“Also, these results are applicable to only this one product, namely SPARK. While the results are not directly transferable to other settings, understanding our evaluative approach and iterative design may be useful to other researchers wishing to create similar products for the pediatric population.”

26. From my point of view, most of the figures are useful. But the readers may be curious about the detailed contents of SPARK. Please provide additional images with better understanding. For example, system architecture, service flow diagram etc. Figure 4 contains poor quality of texts and lines. Please provide a replacement image with better quality.

Response: Please see response to Comment #2 regarding the additional submitted figure. A better quality image for Figure 4 has been submitted.
Editors:

27. All reviewers were enthusiastic about this work, and identified strengths in this manuscript. However, the paper, as is, is focused on the development of the app, and should highlight more clearly what it adds to the field. It would benefit to be better connected to the literature on health information system development and evaluation (see for example Brender et al 2013 STARE-HI – Statement on Reporting of Evaluation Studies in Health Informatics), to clearly highlight the contributions of this paper.

For example, was this a new co-design method? Or specifically adapted to pediatric patients? Was there an evaluation of the system? What was evaluated? Usability? Usefulness? Satisfaction?

Response: Thank you for your feedback. Please see response to Comments #16 and #17 to see how these issues have been addressed.

28. The introduction and discussion sections needs substantial work to connect with the appropriate literature, as suggested by reviewers 2 and 3. Also, given the context of this work (children as participants), details about recruitment and consent would be interesting. Were parents and/or family members involved and if so, how?

Response: Please see response to Comments #16 and #17. This manuscript only describes enrolment of children. The parent and family member portals of SPARK are currently under development and thus, data about their development are not yet available.

29. Also, comments about methods, in particular data collection and analysis, need to be clearly addressed to help the reader assess the reliability and validity of the methods. It remains unclear if the main source of information is the user-patient, or the expert evaluating the experience of patients. For example, were there any video and/or audio recordings? How was the analysis of the observation sessions performed?

Response: To make these items clearer, we have reworded our Methods as follows:

“Understandability was rated externally by experts (two interviewers) while usability, ease of use and usefulness were rated by the user-patient (children themselves).”

There were no videos or audio-recordings made. We have highlighted this by the following addition to the Methods:
“No videos or audio-recordings were made of interviews.”

The analysis of the observation sessions was described in Methods as follows:

“We reported the proportion of participants who were correct or completely correct... “

REFERENCES


