Author’s response to reviews

Title: Designing clinical practice feedback reports: Three steps illustrated in Veterans Health Affairs long term care facilities and programs

Authors:
Zach Landis-Lewis (zachll@umich.edu)
Jennifer Kononowech (henryjen@med.umich.edu)
Winifred Scott (winifred.scott@va.gov)
Robert Hogikyan (hogikyan@med.umich.edu)
Joan Carpenter (joan.g.carpenter@gmail.com)
V. S. Periyakoil (periyakoil@stanford.edu)
Susan Miller (susan_miller@brown.edu)
Cari Levy (cari.levy@va.gov)
Mary Ersek (ersekm@nursing.upenn.edu)
Anne Sales (salesann@umich.edu)

Version: 1 Date: 10 Sep 2019

Author’s response to reviews:

→ Reviewer #1: This paper presents a new method for undertaking user-centred design (UCD) for feedback reports. It presents an example of how it used in the Veterans Health Administration.

→ It is a well-written paper about an important quality improvement topic that doesn't get enough attention.

→ My first main critique of the paper is that there aren't enough clear examples of how to implement the method. I appreciate that additional file 1 and figures 3 and 4 go some way to doing that, but the detail is hidden and isn't clearly linked to the methods. I would recommend bringing it into the main text, perhaps in a table, with summaries of the findings from how it was done in the VA.

Thank you for these critiques and recommendations. We have revised the manuscript to emphasize the primary contribution and to illustrate this in a new table (Table 2) with examples of how the method
was applied in the VA.

→ And, I may be mistaken, but there don't seem to be any usability test findings provided from the VA example. These should be added to the summary table. Figures 3 and 4 are useful, but for example, it would be good to have some description about what their main design differences are, and why those changes were implemented based on their usability testing results.

We have added examples of key findings from usability testing to Table 2 under the column "Observe user". Table 2 includes examples illustrating how each usability finding relates to changes in measures, data and displays in Figures 3, 4, and 5.

→ My other main concern is that we don't have a measure of the impact of this new UCD method. How do we know whether it's worth adopting as feedback designers? This could potentially be demonstrated by improvements in usability metrics over the iterations (which may be solved by addressing my main point above) - but they aren't provided - or some measure of uptake vs non-UCD feedback reports in practice.

This work was part of a larger study that includes a quantitative assessment of the effect of the feedback reports, and reference the protocol paper. We revised the following sentence to emphasize this works relationship to the larger study on page 9: “We illustrate the proposed method through its application in a UCD study to design feedback reports, as part of a broader, large-scale audit and feedback intervention study [16].” However, due to the emergent nature of the development of this method, which we note in the limitations section, we did not plan to assess UCD-derived reports vs. other methods of designing feedback reports. We have added the following sentences to the limitations section on page 19: "We did not quantitatively assess the impact of using the proposed method. Further evaluation is needed to understand the impact of the proposed method on feedback report usability and uptake.”

→ Some more minor comments:
- Is the example about paper or computerised feedback? Needs to be clearer

We have added the phrase "which could potentially be used in electronic or paper form" to the second sentence in the section "Refine display" for the example on page 14.

- There are way too many acronyms - many are used once or twice only and therefore can be just spelled out

We have revised the manuscript to significantly reduce the use of acronyms.

- 'Feedback' and 'A&F' is often used interchangeably. I'd suggest picking one and sticking with it for consistency

We have revised the manuscript to consistently use the term "feedback" instead of "A&F".

- Are any theories / design guidelines / informing usability evidence recommended to design feedback reports as part of the UCD method (i.e. the Display part)? Or were any used in the example
given? I don't think any were mentioned.

We have cited design methods, guidelines, usability studies, and theories throughout the manuscript, but have not recommended specific approaches to designing the manuscript because the method is intended to reflect a general design cycle, rather than specific design techniques. Toward this end we have moved details about UCD techniques into additional files.

- The 'iterations' need more detail - was each iteration usability tested? If so, how many people were used in the tests? How do you know when to stop doing more iterations?

On page 12 we have added the following sentence to provide detail about iterations in the section "Subsequent cycles" in the results: "Each iteration involved the usability testing of the prototype with multiple users and resulted in a refined understanding of users and changes to the prototype reports."

On page 8 we have revised the section of the manuscript "Exiting the UCD cycle" to more clearly indicate how to know when to stop doing more iterations.

- How do you do usability testing over the phone (p18 line 17)? This seems very difficult as is suggested in the discussion section - and there are accepted ways of doing remote usability testing https://www.nngroup.com/articles/unmoderated-user-testing-tools/

Page 14: We have revised the section "Observe user" to clarify our usability testing method, rationale and to cite a source with more details for a qualitative approach to usability testing. On page 19 we have also added further discussion of phone-based usability testing in the limitations section.

- The use of the word 'methods' line 6 p14 is confusing as it's in the results section

We have revised this heading to now read "Application of the method" on page 10.

- A balanced discussion of the pros and cons of this approach would be helpful including other options available to feedback designers instead of this one. For example, Borycki et al ('Usability Methods for Ensuring Health Information Technology Safety: Evidence-Based Approaches. Contribution of the IMIA Working Group Health Informatics for Patient Safety') recommend lab-based, 'in situ'/simulated, and naturalistic testing

Page 19: We have revised the following paragraph to the limitations section and incorporated the recommended citation in an effort to balance our discussion of this approach: “In our application of the proposed method, phone-based usability testing reduced the information we could perceive from participants’ facial expressions and body language that might inform emotional aspects of participants’ comments. Usability testing in an in situ or naturalistic work environment setting for healthcare professionals is likely to support better fidelity to the cognitive processes of perception and comprehension resulting from interaction with reports[28,36]. Phone-based testing was perhaps more feasible because the prototypes we tested were single-page, non-interactive PDF documents that did not require navigation of interface controls or a sequence of actions to perform, which reduced participant’s task complexity.”

- Table 1 - don't you also need someone to conduct the usability testing? Or is that the manager? I would have thought that's a separate job?
We have revised the description of the design team roles, now moved to Additional File 2, to clarify that in our application of the method, the design team lead also conducted usability testing.

- When refining the Measures and Data for the feedback report - isn't it a limitation that you need to work with real-world data to see what the usability problems are, and what information is possible to present in the feedback report? For example, a user may say they would like to see X, Y, Z, but X may not be in the dataset, and it may be impossible to calculate Y or Z.

Thank you for this question. We have revised the first paragraph of the discussion on page 16 to emphasize that the method is used to identify divergent stakeholder perspectives about the performance measures, which guides exploration of what is possible in direct response to expressed user information needs and preferences. We have created Table 2 to demonstrate how refinements to the measures can lead to a search for data items - which may or may not be feasible to use for the purpose of measurement - to create more usable and useful feedback reports. We have also revised the section “Exiting the cycle” on page 15 to reflect that ongoing follow-up has lead to continued revisions to the reports, based on what is appearing in the real-world data.

------------------------------------------

→ Reviewer #2: Thank you for inviting me to review this paper that describes a user-centered design (UCD) process to develop clinical feedback reports. Overall I believe the paper to be well written, of interest and has interesting findings that contribute to the knowledge of developing usable and acceptable clinical feedback reports.

→ Overall I have just a few comments that I believe will improve the readability and understanding of the article. These are of a minor nature only and the methods and description of a UCD process I think is very well described, makes sense (to someone with some experience in designing clinical feedback reports) and sets out a reasonable protocol to follow.

1) I would like to know how much of the process the authors describe in the methods was determined in advance (is there a study protocol they could include?) and how much they refined throughout the process. Were there any steps that were difficult to implement or changes made to the process based upon experience?

Thank you for these comments. We have revised the first paragraph in the limitations section on page 18 to state that the method was developed and clarified as it emerged through application of UCD cycles, and acknowledged the manner in which the method may be specialized to the context of application, as well as a rationale for its generalizability.

2) Page 13 participants - please state the range of people contributing from each site and whether all sites contributed. The following section suggests all sites contributed however it is not clear in this section.

On page 10 we have added the following sentence to this section: "The number of participants from each site ranged from 1 to 4."

3) Page 15 requirements - the authors describe how the UCD process meant that one of the evidence-based key components of feedback was excluded (use of comparators). I would be interested to know more about how the authors/research team came to the decision to exclude this based upon the
feedback received and whether they believe this will potentially impact the report effectiveness. Some
guidance on resolving conflicts and the difficulties of 'single voice' dissenters in UCD would be
helpful. This should also be discussed in the limitations of the study as the UCD process may
recommend some key components are not included in final reports which have a known evidence-base
for increasing effectiveness.

To provide additional information about how the team came to the decision to exclude comparators, we
have added the following sentences to Additional file 2: "A third participant described withholding
feedback reports based on their information content to protect team members from burnout and to
maintain morale. Given this information in the context of broader sensitivity to performance
measurement in the VA, the project team decided to deliver feedback without comparators to increase
the likelihood that feedback reports would be delivered to teams at all, and hence to increase their
effectiveness."

On page 19 we added the following sentences to the limitations section to address the issue of
including items that are not consistent with the known evidence-base: "In applying UCD to the
development of the feedback reports, our identification of divergent stakeholder perspectives resulted
in design choices that reflected the trade-offs we encountered in terms of technical feasibility, available
resources, stakeholder interests, best practice guidance, and evidence about audit and feedback. We
anticipate that UCD methods hold potential to increase the effectiveness of reports, but recognize that
this is an area in which evidence is lacking, and hence a potentially important area for future research."

4) Finally I recommend the authors link their findings to the recent paper by Grimshaw et al 2019 paper
and discuss how the UCD process fits in to the development of implementation laboratories where
head-to-head trials of different ways of providing A&F can advance effectiveness and knowledge
in A&F.

Thank you, we have added the following sentences to relate our study to the generation of evidence in
the A&F Metalab on page 18: "Finally, an implication for the broader adoption of UCD in
feedback report design would require better description of feedback report content and form, to enable
improved evidence generation across research networks, such as the Audit and Feedback Metalab [34].
We recommend that work to better describe and define the components of feedback reports and their
relationships be prioritized to enable better learning from feedback interventions that have been
designed for specific user populations."

→ I hope these comments are useful and constructive. Thank you for the opportunity to review this paper.

→ Reviewer #3: Thank you for the opportunity to review this paper. The authors applied User-Centered
Design to audit and feedback interventions with specific operationalizations that they offer for use by
others in the audit and feedback field.

→ Major comments:

1. The authors wrote that they "extend" step 2 of Witteman and colleagues' framework of user-centered
design, "Develop/refine prototype," by adding steps related to "requirements", "measures", "data", and
"display." This is not an extension; it is an operationalization of this step for audit and feedback. This is
a potentially useful contribution to the field of audit and feedback, but why focus only on
operationalizing step 2? Operationalization of user-centered design steps to audit and feedback could be
applied to the whole cycle of the user-centered design process.

Thank you for your comments. We have revised the manuscript to reconceptualize the method as an operationalization rather than an extension of the UCD cycle. On page 20 we have added the following sentence in the limitations section to call attention to the need to operationalize UCD methods in other steps of the cycle in future research: "We describe steps for operationalizing the “develop/refine prototype” step of the UCD cycle for feedback report design. Future research could additionally explore the operationalization of the UCD cycle for other steps of the cycle, which we have illustrated with examples."

1b. Related to the above, the authors may wish to note that Witteman and colleagues' synthesis specifically omitted the requirements specification stage in other frameworks (namely ISO 9241-210: Human centred design for interactive systems) to allow for the inclusion of co-design methods. The authors may wish to (re-)consult the ISO reference for further information.

We have reconsidered the necessity of requirements specification, and revised the proposed method to leave out the development of requirements as a necessary step. We revised the manuscript, including changing the title and Figure 1, to reflect that there are three necessary steps rather than four. On page 6 we have indicated the option to use explicit requirements in the following sentences under the section, "Develop/refine prototype": “Prototype feedback reports are sketches, drafts, or models of reports that typically contain artificial data and which provide enough information for a user to understand the proposed form and content of the report. Prototypes can be co-created with users or developed without user involvement prior to testing. Refinement of prototypes can be based on user feedback and suggestions during co-creation, or based on requirements that explicitly outline constraints for the report. For example, requirements can be expressed as statements such as “reports must be printable in black and white”, or can be expressed in the form of “user stories” that link a specific user role with a report characteristic and the purpose of the specific requirement [21,22].”

2. Exiting user-centered design cycle does not "ideally happen only after there are no issues identified that would change the major structure of the report." Exiting is about the design solution satisfying user requirements in terms of usability (executing the tasks when using the product but also answering user objectives that can include perception and emotions tied to the product). This may indeed occur when there are no more issues identified that would change the major structure of the report.

Thank you, we have revised this paragraph on page 8 to read as follows “Exiting the UCD cycle can happen when users have demonstrated that the feedback report is usable and that all of the identified user requirements are satisfied, including those relating to users' perceptions and emotional responses to the report. Reaching this point enables the design team to confidently proceed to develop reporting tools or to deliver a design specification for report development.”

3. Page 12: Characteristics of participants in the design of each site are lacking. It would also be good to have site characteristics (such as number of health professionals, number of patients or any other site specifics that may explain differences between requirements and subsequent design.)

We have added Table 1 to describe sites in terms of the number of healthcare professionals and number of patients in Community Living Centers and Home-Based Primary Care programs. On page 19 we have added the following sentence to the limitations section to acknowledge a lack of participant characteristics: “We did not capture participant demographic data or assess participant diversity. We experienced challenges in arranging face-to-face interactions with extremely busy clinicians, and
coordinating schedules in this multi-site, large scale project.”

3b. Related to the above, on page 13 "We compared and contrasted characteristics of site contexts and participants to form generalizable traits of users across sites and to identify contextual differences that reports would need to accommodate." Could the authors describe the differences between sites?

On page 11 we have added the following sentence to indicate that the key differences between sites: "The key differences that reports were adapted to accommodate were the intent to disseminate feedback reports and interest in facility-level comparison, which varied across sites (Additional File 2)." We have described the context and these differences in further detail in additional file 2, under the section "Example 1: Understanding users”.

4. The authors write on page 17: "At the conclusion of the design process we began delivering practice feedback reports to facilities at the 4 demonstration sites and following up with calls to these sites to request additional feedback that would allow for minor improvements to the report design.” Using user-centered design typically implies longer-term surveillance during a specific period of time to check performance of the product, assess whether user needs and requirements were satisfied, and detect problems that may not have been detectable without such follow-up. Was this done in this study? Particularly given the submission to Implementation Science, it would be useful for readers to better understand the authors' plans for continued surveillance post-implementation.

On page 15 we have added the following sentences: "In the two years following report implementation we routinely requested site feedback about the usefulness of the reports, but have not conducted further usability testing. Multiple responses have led to ongoing minor changes and additions to the report, such as shifting the report timing from quarterly to monthly, and adding the provision of patient list data to complement the report data. We plan to provide ongoing maintenance support and have maintained software for the reporting tool in a public code repository (https://github.com/Display-Lab/goals-of-care).”

5. On page 20, the authors write: "Ongoing feedback report adaptation incurs a cost that has not been evaluated, and may not be feasible in smaller-scale interventions that cannot afford to hire a design team. We note that the method is feasible to implement on a small scale, with a single person conducting design team activities, for example iterating prototypes in a spreadsheet.” The authors may wish to reflect on whether a design team with experts is truly necessary. Involving a few users and observing them interacting with the report is one approach that does not typically require a lot of money.

On page 19 we have revised the statement in the limitations section to read as follows: "We did not assess the cost of using the method, in particular the use of a design team with implementation science experts. The method’s value could perhaps be realized at low cost by a single individual, rather than involving a design team.”

6. Step 2. Develop/refine prototype: Figures 3 and 4 present the first and last version of the prototypes. Could the authors present the 12 prototypes (possibly in appendix) and explain the differences between them (ideally within the main manuscript)? It would help understand the results.

We have added Figure 5 showing the current version of the report, and have revised the description of the prototypes in Figures 3 and 4, and we added examples of these refinements to measures and data with explanation of differences in the newly added Table 2. We believe these revisions significantly
improve the description of how changes occurred in successive prototypes, and will help readers to understand the results.

Minor comments:

7. Overall, the heavy use of acronyms makes the manuscript more difficult to follow and less readable than it could be. The authors may wish to consider avoiding acronyms except for key acronyms in the study (e.g., UCD).

We have revised the manuscript to significantly reduce the use of acronyms.

Background:
8. Page 4: The authors write: "Our objective is to describe such a process that builds on prior studies and focuses on the context of large-scale A&F interventions." To which prior studies are they referring? Could they describe them a little, telling the reader what was needed and what they expect to bring to the field? This would help better put their research in context correctly.

We have added description to contextualize these studies in the background section, specifically Colquhoun et al 2017 and Brown et al 2016.

9. Page 4: The authors mentioned, "The purpose of using feedback reports in this context is to promote adoption of new practices and to identify opportunities for improvement." Did they measure any outcomes regarding adoption of new practices after reports were adopted in practice? Later sections (including discussion and conclusion) do not seem to discuss it. Having a user-centered design might be expected to be more efficient.

On page 19 we have added the following sentences to the limitations section: “We did not quantitatively assess the impact of using the proposed method. Further evaluation is needed to understand the impact of the proposed method on feedback report usability and uptake.”

Methods:
10. Page 5: Describing Step 1, understand user, the authors note: "Designing a useful and appropriate feedback report requires the designer to have an understanding of the purpose of the report and the people who will use it." Later, the authors comment on contextual factors. These should be mentioned here, as they are part of the framework they reference. Note that the framework synthesizes other literature such as ISO 9241-210 Human centred design for interactive systems, which specifically notes, "Characteristics of users, tasks, equipment, and the environment (physical and social) in which the product is used." The step "understand user" does not solely refer to users' characteristics.

On page 6 we have revised the description of "Understand user" to emphasize that understanding context is part of this step.

Results:
11. Page 11: what does "life-threatening event" refer to? There is a lack of explanation/clarity (possibly partly because of all the acronyms) regarding patient conditions and the program here.

We have revised the description of the patient population to read "Veterans with serious illness" throughout the manuscript.
12. Page 11: "One site was geographically close." Which location?

We have not provided the name of the site so that the individuals who participated in the design process cannot be identified. We hope that the additional site characteristics provided in Table 1 provide sufficient information to understand the settings for this work.

13. Page 13: was there an interview guide for the process of interviewing, "teams to understand the context for GoCCs, professional roles, Veterans' care processes and environments, and the activities involved in conducting GoCCs"?

We did not use an interview guide as part of the initial site visits. The site visits coincided with site visits for the broader study. We have clarified that the study was part of a larger study in the following sentence in the results section on page 9: “We illustrate the proposed method through its application in a UCD study to design feedback reports, as part of a broader, large-scale audit and feedback intervention study [16].”

14. Page 14: The following is somewhat vague: "We used anecdotes from our observations to support several types of refinements in our understanding, including recognizing false assumptions and establishing ranges of user needs and preferences (Additional File 1, Example 1)." It would be helpful for readers' understanding to have more details in the main text.

We have revised the manuscript to add examples of these refinements in Table 2.

15. Step 2. Develop/refine prototype: It would be helpful to have examples of quotes from field interviews embedded in the text to better understand the team discussions and refinements needed.

We captured field notes but did not record audio from interviews that would allow direct quotes.

Discussion:
16. Did the authors consider conducting formal evaluations (e.g., a factorial experiment) to test some design aspects while implementing in different sites? Why or why not? Could they comment on this?

We did not consider using a factorial trial because of the limited number of sites (4) that had access to the progress note and order set during the initial design phase of the reporting tool. As we indicate in the revised limitations section on page 18, this method emerged as we revised the reports, which prevented us from conceiving of testing design aspects while preparing for the broader trial.