Author's response to reviews

**Title:** The use of mobile phones as a data collection tool: A report from a household survey in South Africa

**Authors:**

Mark Tomlinson ([markt@sun.ac.za](mailto:markt@sun.ac.za))
Wesley Solomon ([fullcirclev@gmail.com](mailto:fullcirclev@gmail.com))
Yages Singh ([Yages.Singh@mrc.ac.za](mailto:Yages.Singh@mrc.ac.za))
Tanya Doherty ([Tanya.Doherty@mrc.ac.za](mailto:Tanya.Doherty@mrc.ac.za))
Mickey Chopra ([Mickey.Chopra@mrc.ac.za](mailto:Mickey.Chopra@mrc.ac.za))
Petrida Ijumba ([Petrida.Ijumba@mrc.ac.za](mailto:Petrida.Ijumba@mrc.ac.za))
Alexander C Tsai ([atsai@post.harvard.edu](mailto:atsai@post.harvard.edu))
Debra Jackson ([djackson@uwc.ac.za](mailto:djackson@uwc.ac.za))

**Version:** 2  **Date:** 20 August 2009

**Author's response to reviews:** see over
Dear Editor,

Many thanks for offering our paper further consideration after the comments from the reviewers. We are grateful for these comments, as they helped to substantially improve our paper. Here is how we responded to their comments, point by point:

Reviewer 1

1. *The current literature review bypasses a thorough review of the application of mobile phones to consumer health informatics/public health research by introducing a few examples and then, explaining there is a dearth of literature describing mobile device applications in low and middle income nations. In short, the authors claim there is no comparative context for the current study. Alternatively, I suggest the manuscript would be enhanced by the provision of a context. I suggest the authors introduce other studies that are similar to this one (regardless of geographical location) and especially introduce any studies that collect survey data, or collect information that is facilitated or consumer-generated rather than monitored automatically. A discussion of similarities and differences also provides an opportunity for the authors to suggest the reasons why the current findings are consistent or inconsistent with prior findings. The comparison additionally provides an opportunity for the authors to discuss how the current study (and its environs) represents specific contributions to the past literature.*

We agree with the reviewer on this point. We have added a more extensive discussion of the context and the use of mobile phones for other purposes.

2. *In terms of methods, it is not clear (until p. 8 of the current manuscript) if community health workers encouraged potential interviewees to take part in the survey, or actually conducted telephone interviews with selected persons. I cannot find an explanation if community health workers entered the results as interviewees responded.*

We have moved the section previously title Training (now titled Data collection and training) to just after the section Setting in the Methods section. We have elaborated how the data collection took place in the homes of respondents face to face.

3. *Assuming the latter assumptions are accurate, other pertinent methodological issues should be mentioned. For example, while the study suggests there was an interviewer/interviewee protocol that was part of community health worker training, the protocol is not explained. In addition, how is the study’s protocol similar or different than the protocol given to a research assistant in standard, by paper, distribution in professionally normative, survey research data collection? How did community health workers contact and select households in the study’s sample? Were the selected households a random sample, or the universe of the*
selected areas? How did the data collectors ensure that only one adult per household served as a respondent?

In the new section (Data collection and training) we elaborate in much more detail how households were selected (the listing process), what the criteria were, as well as what the training comprised.

4. The authors suggest the primary challenge in conducting consumer health informatics/public health survey research is the administrative oversight of problems often associated with data collectors, or employed interviewers. While this is a challenge to statistical reliability and validity, there are other parallel, important considerations (regardless of locale) to ensure the reliability and validity of survey-generated population research. Other challenges to reliability and validity (that are of parallel importance and more often cited) include: access to a domicile, selecting a population sample that is highly similar to overall population characteristics, random sampling of domiciles (dwellings, or a unit of analysis), clustered sampling to reflect population statistics, ensuring the respondent is an adult and only one person is a respondent, question design and bias, readability, and providing reliability and validity measures within an instrument. The authors do not directly address how these issues comparatively are overcome in the current study.

As in point 3 above we elaborate in much more detail in the section Data collection and training how households were selected and what the criteria were. We have also included additional information on training and piloting of the questions used in the survey.

5. The findings suggest it is administratively viable for a trained community health worker to accurately collect and return survey research data by using his/her mobile phone. However, the authors conclusions imply that mobile phone based surveys reduce errors, and curtail high storage and data entry costs. Yet, the evidence seems to indirectly address these variables. Since the study is not comparative, it seems premature to imply that mobile phone based surveys reduce errors, high storage and data entry costs. Instead, the study seems to provide anecdotal findings that there may be comparative costs savings, but these need to be demonstrated via future comparative research.

We agree with Reviewer 1 and their point that we do not have comparative data and therefore are not able to comment on the reduction of errors. We have removed this sentence and we have added a sentence to the limitations section that outlines the fact that we do not have comparative data. We do not agree with the comment that we do not have data on high storage and data entry costs. Data entry costs are zero as is data storage costs.

6. I suggest the authors explain how the training and use of community health workers is superior or equivalent to the use of normally inexpensive research assistants to gather data? What is the likelihood that this technology and approach could be developed so there may be less need for a personal interviewer for data collection?
We have considerably elaborated upon the training section (as outline above). In low and middle income countries with poor literacy levels completing surveys without face to face interviews is not a possibility.

7. I suggest the authors enhance their discussion of the study’s limitations. For example, the current study does not have controls -- the authors do not demonstrate the relative success of their data collection by phone by comparing it with traditional paper-based domicile data collection (or other techniques). While I am not suggesting the authors redo the study, I believe it is important to note that a basis of comparison is missing in the current research and acknowledge the study’s approach is among the study’s limitations. In terms of suggestions for future research, I hope the authors might address how future researchers should conceptually frame and implement comprehensive, comparative research to provide additional evidence that mobile phones are a viable (or superior) method of data collection for consumer health informatics/public health survey research within middle, or low income nations.

We agree with this point of the reviewer and have added to the limitation section a discussion about the lack of a comparison group.

Reviewer 2

1. There are several publications that report the use of mobile phones (and other mobile devices) as a data collection tool. This article should be reviewed and re-submitted as a research article rather than as a technical advance.

We disagree with Reviewer 2 on this point that this is not an innovation. There are no published studies on the use of mobile phones together with a web-based real time monitoring system in such a large survey. There is also no published studies on the use of community health workers completing such a large survey (over 30 000 households) using mobile phones in low and middle income countries. We are also supported in this claim by Reviewer 1 who states that “I am not aware of a comparable project in a parallel setting. So, the manuscript’s primary findings are innovative”. We would therefore argue quite strongly that this is a technical innovation.

2. On page 3, last paragraph; the paper states that “mobile phone technology has been used as a healthcare intervention for chronic, non-communicable diseases;” but reference 5 cited is related to fixed phones, not mobile phones. In addition, reference 7 is related to PDAs, not mobile phones. There are many other reasons of the use of mobile phones for public health purposes. Authors might review and cite examples of mobile health applications from the following report: Vital Wave Consulting. “mHealth for Development”: The Opportunity of Mobile Technology for Healthcare in the Developing World. United Nations Foundation, Vodafone Foundation. February 2009. http://www.vitalwaveconsulting.com/pdf/mHealth.pdf

We agree with the reviewer on these points. We have removed references 5 and 7 and have added new references. We have also added a reference to this report.
3. On page 4, second paragraph; the paper states that “to the best of our knowledge the use of mobile phones as a data collection tool in low and middle countries has not been described.” I cannot agree with this statement. Authors might refer to several papers related to this issue. For example, authors might review and cite: Curioso WH, Karras BT, Campos PE, Buendia C, Holmes KK, Kimball AM. Design and Implementation of Cell PREVEN: A Real-Time Surveillance System for Adverse Events Using Cell Phones in Peru. AMIA Annu Symp Proc 2005; 176-180. http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=16779025 In addition, authors might review and cite some examples from the “mHealth for Development” report.

As outlined in point 1 for this reviewer we disagree with the reviewer on this point. We have cited the reference mentioned as we believe this adds value to the paper.

4. Figure 1. Some screenshots can not be read very well. The authors could improve the screenshots.

We have replaced the unclear screen shots with ones that are much clearer.

5. Mean age of community health workers (CHW) is reported; therefore, standard deviation should be reported. In order to help the reader understand the context, it would help to have other details regarding mobile phone use (e.g. SMS, phone calls), if available: How many CHW were regular SMS users? Do older adults use it much or is it primarily by young adults? All these quantitative data should be reported under results.

We have removed the sentence about the mean age of the community health workers as we do not feel that this is necessary for the message of the paper. We have also added details about sms proficiency in the section on training.

6. Second paragraph: How the authors define an inconsistency? How many inconsistencies were detected? Authors should include a description about inconsistencies in methods.

There is a sentence in the paper on enforced validation which picks up inconsistencies.

7. On page 8, second paragraph; the paper states that “we were able to detect an instance of data fabrication on the day that it occurred.” It’s not clear for the reader if there was only one data fabrication or more. Authors should clarify this issue.

This was the only instance of data fabrication. As a result of the discovery of this instance we met with all the community health workers to explain what had happened. We have added a sentence about this to the paper (see pg 9).

8. Discussion On page 9, first paragraph, the paper states that “This is a significant cost savings.” This might be true, but not data is reflected in the results. On page 9, last paragraph, the paper states that “the Mobile Researcher system is a significantly cheaper option.” This might be true, but not data to support this statement is presented in the results.
We agree with this statement. We have added in a section on the costs of the study but do not make any definitive statement on cost effectiveness because of the lack of a comparison group but we make some tentative conclusions. As already stated, in the limitations section we speak about not having a comparison group.

9. **On page 10, third paragraph, the paper states that “Mobile phones (like PDA’s).” Authors should note that not all PDA are mobile phones.**

The reviewer has misunderstood what is being said in this sentence. This is a statement about the similarities between mobile phones and PDA’s. We are not saying that all PDA’s are mobile phones.

10. **There is no discussion of other mobile health devices already implemented in other settings. Authors should discuss other examples of mobile health projects in developing countries.**

We have added a new section in the introduction on literature from other parts of the world.

11. **Authors should discuss advantages/disadvantages of using proprietary vs. open source software, especially in developing countries.**

We have added paragraph to the end of the discussion section about this.

12. **Some sentences in the conclusion are more suitable for the discussion.**

We have modified the conclusion section and in so doing have responded to this.

Reviewer 3

1. **I would have preferred to have seen a more information on the costs of doing data collection in this manner, since it cites a study that uses PDAs and notes that their approach is less costly. There is no cost data included to allow other researchers to compare relative costs. Since the project didn’t do a direct comparison among paper and pencil, PDA, and cell phone, having some cost data would add a dimension to the report that is lacking.**

We have added a paragraph on this

2. **There was a very little said about problems that were encountered beyond those of assessing data validity, e.g., the way in which the technology offers insights into the reliability of workers in the field.**

We tend to disagree with this comment as there is a great deal stated about the supervisory advantages, the detection of data fabrication.

3. **Data collected in the study was encrypted but no description of encryption methods was included.**
Information on the encryption methods used has been included

4. **On page 7, survey results consist of one line reporting the number of interviews.** After that sentence, “Ne” is written. This needs to be removed. Similarly, in the listing of authors and conflicts of interests, “She” is often spelled “Se”.

We have made the changes.

Yours truly,

Dr Mark Tomlinson