Reviewer's report

Title: A Combination SMS and Transportation Reimbursement Intervention versus Standard of Care to Improve HIV Care Following Abnormal CD4 Test Results in rural Uganda

Version: 1 Date: 29 April 2015

Reviewer: Larry Chang

Reviewer's report:

This is an interesting article on a practical yet innovative intervention to improve HIV care in a low-income setting. It is a very good example of implementation science research, highlighting the many challenges and nuances of conducting this type of study. I commend the authors on the thorough mixed methods and collaborative formative research they conducted to design the intervention.

There are a number of issues which, if addressed, would substantially improve the paper. A main point of confusion is whether the primary intent of the study is to make a pre/post comparison OR to compare the 3 different text messaging strategies (i.e. the randomized component). If both are “primary” or perhaps better described as “key” or something to that effect, then I think the abstract and paper should better reflect this. See below for further comments.

Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)

1. Introduction: 3 para, 1 sentence-Wording is off

2. Methods: I believe it is spelled Dimagi.

Major Compulsory Revisions (while I would classify these as Major, if the authors have sound reasoning for disagreeing with comments, I would find this acceptable)

1. Title: I found the title a bit misleading. It gives the impression that this is an RCT comparing the intervention to SOC and that two groups are occurring in parallel, when this comparison actually uses a pre/post design. I would suggest deleted “versus Standard of Care” or stating explicitly the main study design.

2. Abstract: A general issue with this study, which is exemplified in the results section in the Abstract, is how the 3 intervention groups are handled, i.e. whether to lump or disaggregate these 3 groups. The current strategy in the abstract is to disaggregate but I think the results section should at least report whether any differences were seen between the 3 randomized groups. It is also not always clear throughout the paper when the authors are reporting a lumped comparison vs. disaggregated.

3. Methods: Why were clinicians allowed to select the abnormal CD4 results
threshold? The reasoning behind the use of this criteria was not clear. It may effect generalizability so further insights into this unusual eligibility criteria would be informative, i.e. what exactly would an “abnormally low result”. Discussing how this might impact on generalizability in the discussion would also be helpful.

4. In the methods section, it would be helpful to explain the reasoning behind the three SMS options rather than just what they are.

5. Methods: It would also be good to be clear that your outcome is from test result notification, rather than when the test result was actually performed, e.g. in the Sample Size section one might want to say laboratory result notification rather than just laboratory result.

6. Results: 2nd para-the initial results are presented with the randomized groups disaggregate. I would suggest also providing results with them lumped. Also, as noted above, this results section should preferably report whether there were any differences seen between the 3 intervention groups.

7. Results: This paper would benefit from reporting several pieces of process results. For example, do you have data on time from the actual testing to time of return, i.e. was there a significant delay from the time blood was drawn to when the notification went out? Also, and perhaps most importantly, there are no process results on whether text messages were actually received and/or understood. One would like to know, for example, whether the PIN messages were all or mostly unlocked appropriately. As an extreme example, none of the participants might have received any of the text messages and this is actually only a transport reimbursement intervention.

8. Results: The control group is actually substantially older than the intervention groups and more ART-naïve. Specified threshold is also different. While these may be controlled for in subsequently analyses, they do indicate the possibility residual confounding. This might be something to address in study limitations in the discussion section.

9. Results: Table 1- It is not clear from the table what groups are being compared to produce the p value. Suggest adding an additional column which groups the characteristics for all intervention groups together in one column.

10. Results: Table 2-Also, not clear where the p-value is coming from. Would also add a column and lump the intervention groups.

11. Results: Any data on participant distance from clinic? This would be of interest in regards to generalizability.

12. Discussion: 1st para-again might want to say “abnormal CD4 count notification” if that is what is really going on.

13. Discussion: 2nd para-while this may improve “perceptions of quality of care” you do not have an evidence to this effect. Would modify or delete this statement.
14. Discussion: 2nd para-The fact that those who had normal results also had improved return times may not be because of the SMS messages. They may actually be indicative of a secular change. This should probably be recognized and perhaps included in the limitations section.

15. Discussion: 3rd para-I am not sure this is the “first patient-centered, Health intervention demonstrated to improve patient-provider clinical communication”. Did the WelTel study not do this? In the study limitations, I would probably also discuss the risk of bias given the study was unblinded.

16. Trial Registration-Finally, I will note that this study is registered in Clinicaltrials.gov which is good. However, some of the data from clinicaltrials.gov does not exactly match up with the paper. For example, the primary outcome on the website says receipt and comprehension of text-related message (which do not appear to be reported in this manuscript) rather than time to return to clinic and time to art initiation (outcomes which I could not find on the website).

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.