PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Study protocol and implementation details for a pragmatic, stepped-wedge cluster randomized trial of a digital adherence technology to facilitate tuberculosis treatment completion
AUTHORS	Crowder, Rebecca; Kityamuwesi, Alex; Kiwanuka, Noah; Lamunu, Maureen; Namale, Catherine; Tinka, Lynn; Nakate, Agnes; Ggita, Joseph; Turimumahoro, Patricia; Babirye, Diana; Oyuku, Denis; Berger, Christopher; Tucker, Austin; Patel, Devika; Sammann, Amanda; Dowdy, DW; Stavia, Turyahabwe; Cattamanchi, A; Katamba, Achilles

VERSION 1 – REVIEW

REVIEWER	Shitong Huan Bill & Melinda Gates Foundation Beijing Representative Office, China
REVIEW RETURNED	10-Jun-2020

GENERAL COMMENTS	1. The case enrolment criteria are not clear and may cause the patients under routine management and 99DOTS are different in some characteristics. Eg. Patients under 99DOTS management need to have a cell phone and be literate to read SMS. 2. The patients' adherence information is assessed based on patients' self-report. It is not reliable. How to address this issue, espcially for control groups? 3. In previous 99DOTS pilots, one common issue is the low response rate due to patients' fatigue. How to address this issue in your study? 4. Though IRB committee waivered patients inform concent, I have some concerns on it.

REVIEWER	Rashmi Rodrigues St. John's Medical College, St. John's National Academy of Health Sciences, Bangalore, India I was involved in developing the 99DOT intervention and conducted its first pilot in May 2013. I, however, have no association whatsoever with the others on the team and with Everwell Health Solutions.
REVIEW RETURNED	15-Jul-2020

GENERAL COMMENTS	Please finds comments related to your manuscript below. While I understand that the protocol is already registered and cleared by the relevant Ethics Committees, some clarity regarding certain concepts would be beneficial and has been indicated as such in the assessment.

	Overall, I congratulate you for your effort in this endeavor in the much-needed fight to end TB. It seems the intervention has come a long way from when we started - numbers written on the transparent part of an ATT blister pack and painting them with nail polish to hide them such that they were revealed only when the pills were popped from the blister pack. Best wishes! Abstract: To reflect the changes in the manuscript as necessary. Introduction: Page 4: Line 92: What are the specific cultural and contextual modifications to 99 DOT for implementation in Africa? Page 4: Line 100: Please provide a reference to the PRECEDE framework that you have used. Is it the same as the PRECEDE-PROCEDE framework? Which means only the PRECEDE part of the framework was used. What is the scenario with respect to TB in Uganda- drug susceptible TB, drug resistant TB, levels of adherence, proportion of actually observed currently? The rationale for the study and the use of the intervention along with the current status TB related outcomes in Uganda need to be mentioned here. What proportion of the Ugandan population use mobile phones and what kind of phones do they use? What are the costs of mobile communication in Uganda? What parts of the 99 DOT based intervention strategy did the RE-AIM framework guide and at what stage of development? Conceptual basis for design: While you are testing a public health intervention and aspects of its implementation, type 1 hybrid designs use the individual patient as the unit of randomization. Is it still ok to call it a type 1 hybrid design as in your study clusters are randomized? Please state your hypothesis Methods and analysis: Routine care is usually called- 'standard care'. Study design: Please indicate if this will be a closed cohort, open cohort or continuous recruitment with short exposure. What happens to patients who are recruited when clusters are in the non-intervention phase and move into the intervention phase? Will patients both aspects of the trial? Regarding the step wedge design- while the study is longitudinal and patients receive the intervention during their time in the study the measurement of outcome is at individual level and at one time-point, also this is a rolling cohort, are these considerations used in the sample size estimations? Why was pulmonary TB success rate of <80% and not an overall

	treatment success rate (irrespective of the site of TB) used as an inclusion criterion when you are trying to ensure that the study is pragmatic? If only <2% of the TB units met your inclusion criteria, does it not affect the generalizability of your study findings, making it difficult to scale up the intervention? Study participants: What is the rationale for including only patients with pulmonary TB? What is the diagnostic algorithm followed for pulmonary TB in Uganda? Is NAAT and sputum smear examination done for all patients with presumptive pulmonary TB? What happens if a patient is transferred out after enrollment midway in the study to a TB Unit that has not yet been randomized. What is the perspective that will be used in the cost analysis? If societal, what patient costs will you collect and how? Please indicate if the DOT schedule described in lines 183-190 are as per NTLP guidelines in Uganda. Also, please indicate if the treatment supporter that the patient identifies for DOT is from within the family or a non-related member in the community. Intervention: Please clarify the second component- Daily dosing confirmation using toll free phone calls- do these calls go through or are they what is commonly known as a 'missed call' i.e, call is made but not received? The original intervention worked with a missed call, has this been changed? While it is understandable that patients may choose or not choose to participate in the trial, the ability of the health worker to decide which patient will and will not participate will introduce selection bias, even if using a pragmatic approach. How will you confirm that a dose is actually taken? This, especially, as it is not always necessary that the pills are popped to reveal the number. The number may also be revealed if the flap is pried open. Also, patients may open and discard the pills. What is the masking procedure that will be followed where possible? Wouldn't it be possible to blind the statistician to the intervention? Randomization: It may be better to described randomization in relation to clusters rather than 'groups' such that they can be related to the SW-CRT design used in the study. Data collection and management: Will healthcare workers use their personal phones or the study phone for pictures of registers or both? What proportion of healthcare workers have smart phones with an internet connection? How is data safety and patient confidentiality maintained at all levels? Qualitative: What kind of methods will be used for qualitative data collection? If healthcare workers can administer verbal consent here why can they not for participation in the trial? Costs: Please indicate the perspective from which cost analyses will

	be done. Currently it does not describe the societal perspective mentioned later in the manuscript. Table 2: would be better to describe the primary outcomes first and then the secondary and others. Power and sample size: Please define treatment success and how you will assess it- best described in the outcomes section. What is the current treatment success rate in patients with pulmonary TB in Uganda? Is the 51% success rate for all patients with TB or only patients with pulmonary TB. The calculation of sample size should also indicate how the number of clusters, particioants per cluster and the randomization steps were estimated. Also, if your sample size calculation was based on a power of 90%, why is it not possible to ensure adequate number of clusters and participants to ensure a power of 90%? How is effect of time which is an essential confounder in SW (step wedge)-CRTs considered in your sample size estimation? How are the number of steps involved in Randomization incorporated in the sample size calculation? How does an ICC of 0.001 reflect the number of participants you have included in each cluster? What design effect was used in the calculation of sample size? Also, what is the rationale for assuming an ICC of 0.001, which is very minimal? As this is a SW (step wedge)-RCT please define your cluster. What does a cluster comprise? TB Units or groups of TB Units as indicated in the section on Randomization? Statistical analysis: Statistical analysis is best described in words rather than statistical commands (which can be placed in parenthesis if necessary). How will time be accounted for in the analysis? Why does the model use only fixed effects, especially as outcome is assessed at individual level? Please list the confounding variables that have already been decided a priori- if not here the in the annexures. What outcome paradigms do you propose to use in the sensitivity analysis? Qualitative analysis: What approach will you use analyse the qualitative data? Cost analysis: A societal perspective requires costs from the patients' perspective- which do not seem to be described earlier. Ethics and dissemination: How do you expect patient privacy be compromised through your intervention? Will participants be administered informed consent? The full protocol and statistical code will be made available upon request to whom? Are there any criteria for the investigator to withdraw the patient from the trial or any stop criteria for the trial. Is there any monitoring of adverse events, even though this is not a drug trial? Is there a data safety and monitoring board and will any interim analysis be done? The external monitoring done by Stop-TB cannot

	be considered a part of DSMB. How do you propose to assess intervention fidelity? Discussion: A waiver of consent was obtained for patient-level data collection such that research staff will not be required to be onsite to enroll and consent patients- this is not mentioned under Ethics and dissemination where you say waiver is for obtaining data from registers. Also, isn't it necessary to administer consent to participants prior to requesting them to participate in a research study, where the intervention is not part of the policy? Please discuss the limitations of your study design and of your intervention. Checklist: Please use the Consort Extension for step wedge cluster randomized trials or cluster trials as indicated by the BMJ Open. Both are available. Once again, I wish you all the very best!

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Shitong Huan
Institution and Country: Bill & Melinda Gates Foundation Beijing Representative Office, China
Please state any competing interests or state 'None declared': None declared

1. The case enrolment criteria are not clear and may cause the patients under routine management and 99DOTS are different in some characteristics. Eg. Patients under 99DOTS management need to have a cell phone and be literate to read SMS.

Author response: The enrollment criteria reflect the intention-to-treat (ITT) study population, which includes all adults who initiated treatment for drug susceptible TB at study health centers with the exception of those transferred to other facilities during treatment. The per protocol study population excludes patients during the intervention period who were not enrolled on 99DOTS (due to lack of phone access and other reasons). We present characteristics of patients included in both study populations in Table 1.

2. The patients' adherence information is assessed based on patients' self-report. It is not reliable. How to address this issue, espcially for control groups?

Author response: This trial will compare the 99DOTS-based intervention to routine TB treatment supervision in Uganda. Routine TB treatment supervision (i.e., standard-of-care) involves patients taking their medicines from home (with or without observation by a treatment supporter) and health facility staff assessing adherence via patient self-report at medication refill visits. During both the control and intervention periods, treatment outcomes will be assigned by health facility staff. The reliance on self-reported adherence therefore does not introduce a bias between assessment of treatment outcomes in the control vs. intervention periods.

3. In previous 99DOTS pilots, one common issue is the low response rate due to patients' fatigue. How to address this issue in your study?

Author response: We agree with the reviewer. Therefore, we chose primary and secondary trial outcomes (treatment success, persistence on treatment through the intensive phase, and loss to follow up) that are not based on 99DOTS response rates. Ultimately, the goal is to understand

whether 99DOTS-based treatment supervision improves treatment outcomes as compared to routine care.

4. Though IRB committee waivered patients inform concent, I have some concerns on it.

Author response: As implemented in our study, 99DOTS-based treatment supervision provided the same or greater level of patient oversight as compared to routine care. The institutional review boards at Makerere University School of Public Health and the University of California San Francisco agreed with this assessment, as did the Uganda National Council for Science and Technology. In addition, the data collected to assess patient characteristics and outcomes was the same as that generated as part of routine care. Therefore, the relevant ethical review boards felt criteria for a waiver of informed consent were met.

Reviewer: 2
Reviewer Name: Rashmi Rodrigues
Institution and Country: St. John's Medical College, St. John's National Academy of Health Sciences, Bangalore, India
Please state any competing interests or state ‘None declared’: I was involved in developing the 99DOT intervention and conducted its first pilot in May 2013. I, however, have no association whatsoever with the others on the team and with Everwell Health Solutions.

Dear Authors,

Please finds comments related to your manuscript below. While I understand that the protocol is already registered and cleared by the relevant Ethics Committees, some clarity regarding certain concepts would be beneficial and has been indicated as such in the assessment.

Overall, I congratulate you for your effort in this endeavor in the much-needed fight to end TB. It seems the intervention has come a long way from when we started - numbers written on the transparent part of an ATT blister pack and painting them with nail polish to hide them such that they were revealed only when the pills were popped from the blister pack.

Best wishes!

Comments resulting in changes to the manuscript:

1. Abstract: To reflect the changes in the manuscript as necessary.

Author response: “Hybrid type 1 effectiveness-implementation design” was replaced with “hybrid type 2 effectiveness-implementation design.” Dissemination plans were added to the Ethics and Dissemination section: “Findings will be disseminated through peer-reviewed publications, presentations at scientific conferences and presentations to key stakeholders.”

2. Page 4: Line 100: Please provide a reference to the PRECEDE framework that you have used. Is it the same as the PRECEDE-PROCEDE framework? Which means only the PRECEDE part of the framework was used.

Author response: The reference previously numbered 16 (below) provides background on the framework. We have moved the reference to up to the first mention of the framework. It is now numbered 12.

Green LW, Kreuter MW. Health promotion planning: An educational and ecological approach:

McGraw-Hill; 1993.

Yes, the full name of the framework is the PRECEDE-PROCEED framework. The framework was first developed as the PRECEDE framework in 1974 (acronym for Predisposing, Reinforcing and Enabling Constructs in Educational Diagnosis and Evaluation), and PROCEED was added to the framework in 1992 to capture Policy, Regulatory, and Organizational Constructs in Educational and Environmental Development.

We had referenced PRECEDE only, as this encompasses the aspects of the framework that were employed to design the intervention in our study.

3. The rationale for the study and the use of the intervention along with the current status TB related outcomes in Uganda need to be mentioned here. What proportion of the Ugandan population use mobile phones and what kind of phones do they use? What are the costs of mobile communication in Uganda?

Author response: We have a manuscript describing the pre-trial TB treatment success rates at trial health facilities in press at Journal of Clinical Tuberculosis:
https://doi.org/10.1016/j.jctube.2020.100184
We have added the overall treatment success rate in 2018 as reported in the WHO Global TB Report. In terms of mobile phone use, we have previously published that 69-75% of patients report having access to a mobile phone. We also added this and the citation to the Study Setting section: “Uganda was chosen as the trial setting due to its high TB burden (200 cases/100,000 in 2018) and low treatment success rate (72% in 2017).(1) Previous studies have found 69-75% of TB patients in Uganda have access to a phone.(21, 22)”

4. Conceptual basis for design: While you are testing a public health intervention and aspects of its implementation, type 1 hybrid designs use the individual patient as the unit of randomization. Is it still ok to call it a type 1 hybrid design as in your study clusters are randomized?

Author response: As far as we are aware, the unit of analysis is not a basis for classification of hybrid effectiveness-implementation trials. However, upon reconsidering this, we have re-classified this trial as a hybrid type 2 effectiveness-implementation trial, as we are studying primarily the clinical effectiveness of different implementation strategies for TB treatment supervision (99DOTS vs DOT) while also assessing their reach, adoption and implementation.

5. Please state your hypothesis

Author response: Under “Study Aims” we have now included the sentence: “Our primary hypothesis is that 99DOTS-based TB treatment supervision will improve TB treatment outcomes. Our secondary hypotheses are that the 99DOTS-based strategy will have high uptake among patients and providers, and be cost-effective.”

6. Study design: Please indicate if this will be a closed cohort, open cohort or continuous recruitment with short exposure. What happens to patients who are recruited when clusters are in the non-intervention phase and move into the intervention phase? Will patients both aspects of the trial?

Author response: We used a repeated cross-sectional design. In each month of trial recruitment, all eligible patients initiated on TB treatment at participating facilities were included. Patient outcomes were assigned to the health facility and study period (control/buffer/intervention) in which the patient initiated treatment. After a facility switched to the intervention, new patients initiating treatment were assigned to the intervention period. Facility staff were instructed to enroll new patients 99DOTS and

to continue supervising patients already on treatment using standard care. In a per protocol analysis, we exclude eligible patients from the control period that were enrolled on 99DOTS and eligible patients from the intervention period who were not enrolled on 99DOTS.

We have added the following sentences to the Study Design section:
• “We included repeated cross-sectional samples of eligible individuals initiating TB treatment at participating health facilities at 8 time points (months). Patient outcomes were assigned to the health facility and month in which they initiated treatment.”
• “Following implementation of the intervention at each site, health facility staff were instructed to offer 99DOTS to all new eligible patients initiating treatment and to continue supervising patients already on treatment using routine care.”

7. If only <2% of the TB units met your inclusion criteria, does it not affect the generalizability of your study findings, making it difficult to scale up the intervention?

Author response: TB treatment is decentralized in Uganda, and most facilities initiate fewer than 10 patients on TB treatment each month. In order to enroll enough patients to assess the effectiveness of the 99DOTS-based intervention, we selected larger facilities. We have added this as a potential limitation in the discussion: “Uptake and effectiveness of 99DOTS may be different at lower volume health centers.”

8. Please indicate if the DOT schedule described in lines 183-190 are as per NTLP guidelines in Uganda. Also, please indicate if the treatment supporter that the patient identifies for DOT is from within the family or a non-related member in the community.

Author response: We have added “as per NTLP guidelines” to the first sentence in the Routine Care section and the phrase, “a family member or non-family community member” to describe treatment supporters. The type of treatment supporter varies by facility and we have described this in a manuscript in press at Journal of Clinical Tuberculosis: https://doi.org/10.1016/j.jctube.2020.100184

9. Intervention: Please clarify the second component- Daily dosing confirmation using toll free phone calls- do these calls go through or are they what is commonly known as a ‘missed call’ i.e, call is made but not received? The original intervention worked with a missed call, has this been changed?

Author response: We have added the following sentence to the Intervention section: “When patients call to confirm dosing, they hear a recorded educational or motivational message about continuing and completing TB treatment.”

10. Randomization: It may be better to described randomization in relation to clusters rather than ‘groups’ such that they can be related to the SW-CRT design used in the study.

Author response: We replaced the word “group” with “allocation sequence” and “block”:
• “The 18 health facilities will be randomly assigned to one of the six allocation sequences (Figure 1) using a simple, unrestricted two-stage process.”
• “First, health facilities (i.e., clusters) will be randomly assigned into six blocks of three by having health facility representatives each draw 1 of 18 balls (3 each labeled A-F) from an opaque bag. Blocks will then be randomly assigned to an intervention initiation time, which will occur at equally spaced one-month intervals during the trial, by drawing of 6 balls labeled 1-6 from an opaque bag.”

11. Costs: Please indicate the perspective from which cost analyses will be done. Currently it does not describe the societal perspective mentioned later in the manuscript.

Author response: We added the following sentence to the Health System Cost Data Collection section: “Costing and cost-effectiveness analyses will focus on the health system perspective”. We have replaced the phrase “societal perspective” in the Analysis section with “health system perspective” to better reflect the approach used.

12. As this is a SW (step wedge)-RCT please define your cluster. What does a cluster comprise? TB Units or groups of TB Units as indicated in the section on Randomization?

Author response: We identify each healthy facility as a cluster. The health facilities (i.e., clusters) were grouped in blocks of 3 and each block was randomly allocated. We have clarified the language describing this in the Randomization section: “First, health facilities (i.e., clusters) will be randomly assigned into six blocks of three by having health facility representatives each draw 1 of 18 balls (3 each labeled A-F) from an opaque bag. Blocks will then be randomly assigned to an intervention initiation time, which will occur at equally spaced one-month intervals during the trial, by drawing of 6 balls labeled 1-6 from an opaque bag.”

13. What is the masking procedure that will be followed where possible? Wouldn’t it be possible to blind the statistician to the intervention?

Author response: We have clarified that both study staff and investigators will be blinded to aggregate TB outcomes by study period.

14. How is data safety and patient confidentiality maintained at all levels?

Author response: In routine care, patient identifying information is entered into paper registers, which are easily accessed and can easily be stolen. For the study, photos of the registers are stored on a secure, password-protected server only accessible to study staff. In addition, patient data are entered into a secure, 99DOTS server for those enrolled on treatment. Again, access to the server is password-protected. We have added the phrases “password-protected” and “only accessible to research staff” to the description of photo storage. These aspects are currently mentioned in the Patient Level Data Collection section: “Project staff will train two health workers at each site (one primary, one backup) identified by the health facility director to take photos of the register every 2-4 weeks for the duration of the project using a camera-enabled smartphone, and to upload the photos to a central secure, password-protected server, only accessible to staff. Health workers will be trained to delete photos from the phone after upload confirmation. Completeness of TB treatment registers will be assessed, and at quarterly site visits, study staff will provide re-training as needed and resolve data cleaning queries. Study staff will extract data from photos of TB registers and enter data into a secure database use Research Electronic Data Capture software (REDCap).(23)”

Statistical analysis:
15. Statistical analysis is best described in words rather than statistical commands (which can be placed in parenthesis if necessary).

Author response: Additional clarification was added to the Analysis section on the structure of the planned analysis models: “The primary effectiveness analysis will be conducted at the health facility level using multivariable mixed effect logit models with random effects for site and fixed effects for trial period, time, and patient-level confounders (using Stata’s melogit and meqrlogit commands).”

16. How will time be accounted for in the analysis? Why does the model use only fixed effects, especially as outcome is assessed at individual level?

Author response: The primary analysis is a mixed effect logit model with random effects for site and

fixed effects for trial period, time (trial month), and patient-level confounders. A clarification was added to the Analysis section: “Models will adjust for the longitudinal design (indicator variable for each trial month) and clustering by site (random effect for health facility).”

17. Please list the confounding variables that have already been decided a priori- if not here the in the annexures.

Author response: We will adjust for sex, HIV status, disease class (bacteriologically confirmed vs. clinically diagnosed), and TB type (new vs. retreatment). This information has been added to the Analysis section: “Potential confounders, selected a priori, including sex, HIV status, disease class (bacteriologically confirmed vs. clinically diagnosed), and TB type (new vs. retreatment), will be included in the model as fixed effects.”

18. Cost analysis: A societal perspective requires costs from the patients' perspective- which do not seem to be described earlier.

Author response: We agree with the reviewer. We have replaced the phrase “societal perspective” in the Analysis section with “health system perspective” to better reflect the approach used.

19. A waiver of consent was obtained for patient-level data collection such that research staff will not be required to be onsite to enroll and consent patients- this is not mentioned under Ethics and dissemination where you say waiver is for obtaining data from registers.

Author response: The phrase “such that research staff will not be required to be onsite to enroll and consent patients” was added to the Ethics and Dissemination section. It now reads, “To ensure that the study captures all eligible adults initiating treatment at study sites, a waiver of informed consent was granted to access routine TB treatment data recorded in standard NTLP registers, such that research staff will not be required to be onsite to enroll and consent patients.”

20. Please discuss the limitations of your study design and of your intervention.

Author response: We have added limitations to the Discussion section as follows:
“Limitations of such a pragmatic trial design include less control over intervention delivery and potential limited uptake of the intervention given the wide inclusion criteria (all adults initiating treatment for drug-suceptible pulmonary TB). In order to enroll enough patients to assess the effectiveness of the 99DOTS-based intervention, we selected facilities that treat larger numbers of TB patients. Uptake and effectiveness of 99DOTS may be different at lower volume health centers.”

21. Checklist: Please use the Consort Extension for step wedge cluster randomized trials or cluster trials as indicated by the BMJ Open. Both are available.

Author response: We have attached the CONSORT checklist for the stepped-wedge cluster randomized trials with updated page numbers following revisions.

Responses to comments for which no changes were made to the manuscript:

22. Introduction: Page 4: Line 92: What are the specific cultural and contextual modifications to 99 DOT for implementation in Africa?

Author response: As shown in Figure 3, we re-designed the 99DOTS envelope to reduce stigma, provide clear and simple instructions to guide pill taking, and include appropriate counseling messages identified by local health workers. We also re-designed the patient call in experience to

involved educational/motivational messages recorded by local health workers instead of a generic audio tone. We have two manuscripts submitted describing the human centered design process used to make the cultural and contextual modifications. The following conference abstracts also describe this work:

https://academyhealth.confex.com/academyhealth/2019di/meetingapp.cgi/Paper/35912

23. What is the scenario with respect to TB in Uganda- drug susceptible TB, drug resistant TB, levels of adherence, proportion of actually observed currently?

Author response: The W|HO estimates that the total incidence of TB in Uganda in 2018 was 86,000 cases, or 200 cases per 100,000 population. Of these, 1,500 cases were rifampicin-resistant (3.5 cases per 100,000 population). TB treatment adherence is not routinely collected or reported. Treatment completion rates in Uganda are generally low (72%) and we have a manuscript in press at Journal of Clinical Tuberculosis describing the baseline treatment success rates at the study facilities: https://doi.org/10.1016/j.jctube.2020.100184

24. What parts of the 99 DOT based intervention strategy did the RE-AIM framework guide and at what stage of development?

Author response: We used the RE-AIM framework as an evaluation rather than an intervention design framework, which is its most common use. It guided the selection of outcomes to assess as shown in Table 2.

25. Methods and analysis: Routine care is usually called- ‘standard care’.

Author response: We defer to the Editors here. We believe specifying routine care is important as it highlights the pragmatic nature of the trial. We did not intervene to ensure that routine care met the expected standard of care.

26. Regarding the step wedge design- while the study is longitudinal and patients receive the intervention during their time in the study the measurement of outcome is at individual level and at one time-point, also this is a rolling cohort, are these considerations used in the sample size estimations?

Author response: Yes. The sample size was calculated assuming a harmonic mean of 15 patients initiated on treatment per month, with the outcomes of those patients assigned to the month in which they initiated treatment.

27. Why was pulmonary TB success rate of <80% and not an overall treatment success rate (irrespective of the site of TB) used as an inclusion criterion when you are trying to ensure that the study is pragmatic?

Author response: We focused on the treatment success rate for pulmonary TB since that is the study population (ie, patients with other forms of TB were excluded).

28. Study participants: What is the rationale for including only patients with pulmonary TB? What is the diagnostic algorithm followed for pulmonary TB in Uganda? Is NAAT and sputum smear examination done for all patients with presumptive pulmonary TB?

Author response: Pulmonary TB is the most common form of TB and the most infectious. We

therefore focused on this population. In addition, the treatment duration is different for different forms of extra-pulmonary TB. The primary method of TB diagnosis at health facilities was either sputum smear microscopy or Xpert MTB/RIF. The study focused on those who were diagnosed with and initiated on treatment for pulmonary TB and did not intervene on the method of diagnosis. In addition, a major criticism of the initial round of TB REACH Wave 5 studies from the STOP TB Partnership Board was that these studies had confusing results because they included all forms of TB (pulmonary/extra-pulmonary, drug susceptible/drug resistant). Therefore, it was a requirement of the funder to focus on a single form of TB and we picked the most common form (drug susceptible pulmonary TB).

29. What happens if a patient is transferred out after enrollment midway in the study to a TB Unit that has not yet been randomized.

Author response: Patients who were transferred out to another facility to complete treatment will be excluded. This is stated under the Study Participants section: “All adult patients treated for drug-susceptible pulmonary TB at participating treatment units during the study period will be eligible for inclusion. Children, patients treated for drug-resistant or extrapulmonary TB, and patients transferred to another facility to complete treatment will be excluded.”

30. What is the perspective that will be used in the cost analysis? If societal, what patient costs will you collect and how?

Author response: Due to funding limitations, we will focus only on assessing the costs to the health system, as now stated in the Health System and Cost Data Collection section.

31. While it is understandable that patients may choose or not choose to participate in the trial, the ability of the health worker to decide which patient will and will not participate will introduce selection bias, even if using a pragmatic approach.

Author response: We agree with the reviewer’s comment. However, in practice, decisions to enroll or not enroll a patient on 99DOTS will be made by health workers and patients. Please note due to the pragematic implementation, this does not reflect a decision to participate in the trial (all eligible patients are included in the ITT analysis, regardless of whether they are enrolled on 99DOTS during the intervention period). This is the reason we specify both an ITT and per protocol analysis.

32. How will you confirm that a dose is actually taken? This, especially, as it is not always necessary that the pills are popped to reveal the number. The number may also be revealed if the flap is pried open. Also, patients may open and discard the pills.

Author response: The issue the reviewer raises applies to both the routine care and intervention periods. The primary and secondary trial outcomes of the trial therefore focus on treatment outcomes as recorded by health workers rather than adherence.

33. Data collection and management: Will healthcare workers use their personal phones or the study phone for pictures of registers or both? What proportion of healthcare workers have smart phones with an internet connection?

Author response: Health care workers are provided smartphones by the study. This is described in the Intervention section: “Health facilities using 99DOTS to manage TB patient treatment will be provided one smartphone per facility, minimal funds to cover cell phone data, and an average of 300,000 USh (approximately $82 USD) per month to facilitate patient follow-up.” The cost of providing phones will be accounted for in the costing analysis.

34. Qualitative: What kind of methods will be used for qualitative data collection? If healthcare workers can administer verbal consent here why can they not for participation in the trial?

Author response: We indicate that research staff, not routine health workers, will administer verbal consent to patients selected for participation in interviews:

“Surveys will be conducted with 10 randomly selected eligible patients enrolled on 99DOTS (5 women and 5 men) and 1-2 eligible providers per site (180 total patients, 18-36 total providers) beginning in Month 10 of the trial to assess the acceptability of the 99DOTS-based intervention strategy. Research staff will contact selected patients by phone to review the verbal consent script, answer questions the patient may have, and administer the survey to consenting patients. Provider surveys and interviews will be conducted during quarterly site visits beginning in Month 13 of the trial to assess factors influencing the adoption and implementation of 99DOTS.”

35. Table 2: would be better to describe the primary outcomes first and then the secondary and others.

Author response: We defer to the Editor. We believe the current structure of the table reflects and demonstrates use of the RE-AIM framework. There is a single primary outcome and all other outcomes are secondary.

36. Power and sample size: Please define treatment success and how you will assess it- best described in the outcomes section.

Author response: In the second sentence of the Outcome section we specify the definition of treatment success: “The primary effectiveness outcome is the proportion of patients who complete TB treatment, defined as having an outcome of “cured” or “treatment completed” recorded in the unit TB treatment register.”

37. What is the current treatment success rate in patients with pulmonary TB in Uganda? Is the 51% success rate for all patients with TB or only patients with pulmonary TB.

Author response: The WHO TB report indicates that the treatment success rate for all new and relapse cases was 72% in 2017. However, in our experience, the treatment success rates are lower when actually reviewing individual facility-level data. We used 51% in the sample size calculation because this was the treatment success rate for drug-susceptible pulmonary TB patients in 2017 at health facilities participating in the trial.

38. The calculation of sample size should also indicate how the number of clusters, particioants per cluster and the randomization steps were estimated. Also, if your sample size calculation was based on a power of 90%, why is it not possible to ensure adequate number of clusters and participants to ensure a power of 90%? How is effect of time which is an essential confounder in SW (step wedge)-CRTs considered in your sample size estimation? How are the number of steps involved in Randomization incorporated in the sample size calculation?

Author response: We used the steppedwedge command in Stata to estimate power and select the number of health facilities and steps to be included. As described in detail here (https://journals.sagepub.com/doi/pdf/10.1177/1536867X1401400208), the command is based on a linear mixed model to describe the data with time as a fixed factor at (t + 1) levels and with a random effect for intercluster variation. We considered different numbers of health centers, average cluster size using data from the health centers, and number of steps. In so doing, we determined that with 18

health facilities, a harmonic mean of 15 patients per step per cluster, 6 steps, 3 sites switching at each step, and an ICC of 0.001, we would have 89% power to detect an 10% difference in the primary outcome between intervention and control periods.

39. How does an ICC of 0.001 reflect the number of participants you have included in each cluster? What design effect was used in the calculation of sample size? Also, what is the rationale for assuming an ICC of 0.001, which is very minimal?

Author response: The number of participants in each cluster was determined using the harmonic mean number of eligible patients initiating TB treatment at participating facilities in 2017 and was estimated to be 15. The intracluster correlation coefficient (ICC) compares the within-group variance with the between-group variance in the outcome of interest and was calculated using pre-trial data from 2017. A major strength of the trial design is that we had pre-trial data from participating health facilities to estimate cluster size and ICC.

40. What outcome paradigms do you propose to use in the sensitivity analysis?

Author response: We are not sure what the reviewer refers to by outcome paradigms.

41. Qualitative analysis: What approach will you use analyse the qualitative data?

Author response: Patient and provider surveys will be analyzed descriptively and interpreted in the relation to the dimensions of RE-AIM, the Theoretical Domains Framework, and the Unified Theory of Acceptance and Use of Technology. Patient and provider interviews will be coded using qualitative software and assessed for themes.

42. Ethics and dissemination: How do you expect patient privacy be compromised through your intervention?

Author response: We do not expect that patient privacy will be compromised beyond the risk of this occurring in routine care. 99DOTS has been scaled up as the primary mode of TB treatment supervision in India, which has the highest TB burden in the world. As with any study, there is a small risk for loss of confidentiality and stigma should personal health information, including HIV or TB status, be disclosed. We have put in place standard precautions to minimize this risk including minimizing access to study databases, storing electronic data on secure and password-protected databases, and using de-identified data sets for analysis.

43. Will participants be administered informed consent?

Author response: No. The trial was approved with a waiver of informed consent by institutional review boards at Makerere University School of Public Health and the University of California San Francisco. Patients and providers participating in study surveys and interviews will be administered informed consent.

44. The full protocol and statistical code will be made available upon request to whom?

Author response: We defer to the editor here on the standard language preferred. We would make these available to anyone involved in health care research or delivery.

45. Are there any criteria for the investigator to withdraw the patient from the trial or any stop criteria for the trial.

Author response: Patients may withdraw from using 99DOTS at any time. There are no investigator-initiated criteria for withdrawing a patient from the trial or 99DOTS.

46. Is there any monitoring of adverse events, even though this is not a drug trial?

Author response: The institutional review boards approved the trial without specific monitoring of adverse events as they agreed the risk of such events was low, and the level of treatment supervision provided by 99DOTS was greater than what is done in routine care.

47. Is there a data safety and monitoring board and will any interim analysis be done? The external monitoring done by Stop-TB cannot be considered a part of DSMB.

Author response: As noted above, due to the minimal risk nature of the study, neither the funder nor the IRBs required a DSMB. The Principal Investigators will be responsible for monitoring the data, assuring protocol compliance, and conducting safety reviews on a quarterly basis. External monitoring is provided by a Stop TB Partnership project officer and Monitoring and Evaluation consultant. The Principal Investigator will submit regular progress reports, including recommendations on whether the project should continue unchanged, require modification, or close to enrollment.

48. How do you propose to assess intervention fidelity?

Author response: Outcomes related to the implementation domain of the RE-AIM framework (Table 2) include measures of fidelity. The proportion of daily SMS sent by 99DOTS and received by patients and the proportion of IVR calls sent by 99DOTS and received by patients (implementation) measure the extent to which 99DOTS was delivered and used as intended.

Discussion:
49. Also, isn't it necessary to administer consent to participants prior to requesting them to participate in a research study, where the intervention is not part of the policy?

Author response: The trial was approved with a waiver of informed consent by institutional review boards at Makerere University School of Public Health and the University of California San Francisco. The IRBs agreed that the trial satisfied criteria for a waiver of informed consent, including minimal risk to patients (99DOTS provides a greater level of treatment supervision including daily adherence monitoring than occurs under routine care) and outcome assessment using data generated as part of routine care.

VERSION 2 – REVIEW

REVIEWER	Rashmi Rodrigues St. John's National Academy of Health Sciences, Bangalore, India I was involved in developing the first version of 99DOT intervention in 2013
REVIEW RETURNED	06-Nov-2020

GENERAL COMMENTS	Thank you for the detailed responses that are very informative and add value and context to the manuscript. Here are a few comments based on your responses: 1. Adherence to treatment is essential to trials involving interventions that focus on improving adherence even though it is not the primary outcome. It is therefore essential to measure adherence uniformly in both arms with uniform methods in addition to adherence measured

	via the intervention in the intervention arm. Also, it would be good to include this issue with measurement of adherence in the study limitations- which is different from intervention fidelity. 2. As per ITT analysis all those randomised are analysed. If for any reason exclusion of anyone randomised is decided a priori in the protocol then it may be considered a modified ITT. 3. Apologies for misphrasing the question on sensitivity analysis- the question seeks to clarify the independent variables that will be used in the sensitivity analysis and their range of uncertainty. 4. The analysis states "excluding patients who are not enrolled on 99DOTS during the intervention period at each site'. However, the response to Q 31 states that "all eligible patients are included in the ITT analysis, regardless of whether they are enrolled on 99DOTS during the intervention period" - please clarify the discrepancy. 5. The limitations of the intervention itself are not sufficiently discussed/ mentioned. I think it is important to state these limitations such as the audience/ readers can weigh the pros and cons of implementing the intervention- as we are all well aware no public health intervention is devoid of limitations. Once again, congratulations and best wishes!

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Rashmi Rodrigues
Institution and Country: St. John's National Academy of Health Sciences, Bangalore, India
Competing interests : I was involved in developing the first version of 99DOT intervention in 2013

1. Adherence to treatment is essential to trials involving interventions that focus on improving adherence even though it is not the primary outcome. It is therefore essential to measure adherence uniformly in both arms with uniform methods in addition to adherence measured via the intervention in the intervention arm. Also, it would be good to include this issue with measurement of adherence in the study limitations- which is different from intervention fidelity.

Author response: We focus on treatment outcome rather than adherence as this is what is reported by National TB Programs. In addition, study procedures required to rigorously assess adherence are likely in and of themselves to impact adherence and retention on treatment. Therefore, we do not verify adherence during either the control or intervention period. Nonetheless, we have added the following to the limitations section of the Discussion:
• “Limitations of such a pragmatic trial design include less control over intervention delivery, potential limited uptake of the intervention given the wide inclusion criteria (all adults initiating treatment for drug-susceptible pulmonary TB) and inability to rigorously verify medication adherence during the control and intervention periods.”

2. As per ITT analysis all those randomised are analysed. If for any reason exclusion of anyone randomised is decided a priori in the protocol then it may be considered a modified ITT.

Author response: The ITT population will include all eligible patients. We have tried to clarify this in the analysis section:

• “Analysis will be done for intention to treat (all eligible patients) and per protocol (excluding patients who are not enrolled on 99DOTS during the intervention period at each site) populations.”

3. Apologies for misphrasing the question on sensitivity analysis- the question seeks to clarify the independent variables that will be used in the sensitivity analysis and their range of uncertainty.

Author response: We have clarified in the analysis section: “To further assess the robustness of our findings, we will conduct sensitivity analyses that include patients who initiated treatment during the buffer period, impute outcomes for patients lost to follow up and compare treatment outcomes using permutation tests.(24)”

4. The analysis states "excluding patients who are not enrolled on 99DOTS during the intervention period at each site'. However, the response to Q 31 states that "all eligible patients are included in the ITT analysis, regardless of whether they are enrolled on 99DOTS during the intervention period" - please clarify the discrepancy.

Author response: Sorry for the confusion. As noted in Response 2, we have clarified as follows:

• “Analysis will be done for intention to treat (all eligible patients) and per protocol (excluding patients who are not enrolled on 99DOTS during the intervention period at each site) populations.”

5. The limitations of the intervention itself are not sufficiently discussed/ mentioned. I think it is important to state these limitations such as the audience/ readers can weigh the pros and cons of implementing the intervention- as we are all well aware no public health intervention is devoid of limitations.

Author response: We agree it is important to state the limitations of an intervention. However, we feel this is best done in a manuscript describing the intervention evaluation results and will certainly include this when we publish the trial results. Here, we present the trial protocol and therefore focus on the strengths and limitations of the study design.