PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Capturing patient-reported and quality of life outcomes with use of shorter regimens for drug-resistant tuberculosis: mixed-methods sub-study protocol, TB PRACTECAL-PRO
AUTHORS	Stringer, Beverley; Lowton, Karen; James, Nicola; Nyang'wa, Bern-Thomas

VERSION 1 – REVIEW

REVIEWER	Metcalfe, John UCSF
REVIEW RETURNED	21-Sep-2020

GENERAL COMMENTS	This is a nice description of a sub-study protocol and I have only a few minor comments. --More detail would be appreciated on how the authors will sample healthy controls at each site/country. --Why did the authors chose SF 12 over another widely accepted, brief QOL survey: WHO Quality of Life-BREF (WHOQOL-BREF)? --Since the authors chose a brief survey version for general QOL (SF 12 over SF 36), could they justify why they chose the relatively long SGRQ over the COPD Assessment Test, a briefer respiratory-specific questionnaire. --Since clofazimine often causes reversible non-melanotic hyperpigmentation that can lead to identifiability and potential stigma (which can vary significantly by ethnogeography), the authors should anticipate stratifying their results by whether clofazimine was part of the treatment regimen or not.

REVIEWER	Hoddinott, Graeme Stellenbosch University Faculty of Medicine and Health Sciences, Desmond Tutu TB Centre
REVIEW RETURNED	02-Oct-2020

GENERAL COMMENTS	Thank you for the opportunity to review. This sub-study protocol is clearly described and the sub-study addresses an important gap and is appropriately designed. Suggestions: 1. The sub-section on pages 3-4 on patient reported outcome measures - this seems unnecessarily long and detailed. There is sufficient recognition of the value of these, and the protocol aim could not be achieved in any other way, therefore this extended justification seems an odd inclusion. I suggest that this be amended to instead address HOW understanding patients' QoL on

	the specific shorter regimen evaluated in PRACTECAL will be useful to implementation, roll-out, and future further iterations of refinements to treatment regiments (i.e., informing development priorities). You can then place greater emphasis on the novelty of an objective measure of QoL impact of the shorter regimen because you have nested this in the trial. 2. Page 4, lines 20-24 - If the SF36 is the most commonly used, why did you then choose the SF12. Doesn't the reduction in items give less detail for describing QoL? Some justification at this point would be useful. 3. Page 4, sub-section on qualitative PRO measures in Tuberculosis - I'm not clear about the point of this paragraph. There is a significantly larger literature reporting qualitative data on the general experiences of people with TB (adults and children). There is a smaller, but still robust, literature reporting on qualitative data on the quality of life of patients with TB. If this is a summary of those literatures, then there are notable exclusions. However, I think the authors are still making the more general methodological point that QoL can be studied with qualitative data. If so, see point 1, this just seems unnecessary. It is more important to summarize the literature on qualitative data about patient experiences of MDR-TB treatment, or how MDR-TB treatment has impacted their quality of life (more narrow). 4. Page 5, line 8 - The quantitative data are exclusively on QoL impact. To describe this as a study about 'perceptions, expectations and experiences' seems too broad, even though there will also be qualitative data collected. Suggest revising to read 'patient quality of life relevant experiences' instead. 5. Secondary objective 2 - why would we assume that the SGRQ and SF-12 DON'T have utility in TB clinical trials? These are well-established measures. Similarly, of course some qualitative data can be collected as part of clinical trials (this has been done before!), and if done well, this should have utility. Perhaps this is a question of feasibility rather than utility? But again, still, that seems to be a question that we already have the answer to. Suggest removing this objective. 6. Page 5, lines 43-47 - this repeats information already stated in the objectives. Suggest removing here. 7. Page 5, sub-section on patient and public involvement - this manuscript is reporting on PRACTECAL-PRO, not the overall trial. Were patients and the public involved in PRACTECAL-PRO? 8. Page 6, sampling - it is unclear why you would not administer these QoL instruments to all PRACTECAL participants. Why sub-sample at all if you intend to do a descriptive analysis? 9. Page 6, healthy controls - I suggest that the point isn't that they are 'healthy', but rather that they experience a QoL that is the average for this population but not people living with TB in this population. I.e., many people would argue that people's average 'healthy' QoL in Scandinavia is higher than these study sites. We're looking for an average person's QoL in this place, not a 'healthy' person's QoL for the comparator, correct?

	10. Page 7, second paragraph under the sub-heading 'Surveys' - this does not seem to be about the instruments but rather a repetition of aim/objectives. Suggest removing. 11. Page 8, data analysis - suggest describing this as things you will describe, rather than 'objectives include:' 12. Page 8, lines 27-33 - repeats from sampling section. Suggest removing. 13. Page 8, qualitative data - commend the authors for a strong methods description here. 14. Page 8, Discussion section - is this relevant in the absence of findings? It can only duplicate the rationale as outlined in the introduction. Suggest removing.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:
--More detail would be appreciated on how the authors will sample healthy controls at each site/country.

Thank you for this comment, more details (in red font) have been added to the manuscript on page 7, line 259, under the 'sampling' section.
e.g. Survey data from 108 healthy controls from the general population in the three study countries will be collected at one timepoint only, matched as closely as possible to the age and sex profile of trial patients. Each site will opportunistically identify participants from the community setting which may include personal contacts and colleagues not working on the PRACTECAL study. The healthy control will be screened for TB symptoms using a symptom screening tool outlined in trial standard operating procedures; only those who screen-negative will participate. For those screen-positive, we will offer further investigation and treatment using established programmatic protocols. Additionally, we will ask each potential healthy control to tell the investigator if they consider themselves generally healthy and with no significant illnesses. Prospective participants reporting any health problems will also be excluded from the sub-study.

--Why did the authors chose SF 12 over another widely accepted, brief QOL survey:
WHO Quality of Life-BREF (WHOQOL-BREF)?

We chose the SF12 as whilst the WHOQOL-BREF may be more suitable for measuring global QOL, we were interested in measuring the impact of interventions on health-related QOL, which indicated the short form series, (Huang, IC., Wu, A.W. & Frangakis, C. Do the SF-36 and WHOQOL-BREF Measure the Same Constructs? Evidence from the Taiwan Population*. Qual Life Res 15, 15–24 (2006). https://doi.org/10.1007/s11136-005-8486-9)
--Since the authors chose a brief survey version for general QOL (SF 12 over SF 36), could they justify why they chose the relatively long SGRQ over the COPD Assessment Test, a briefer respiratory-specific questionnaire.

Thank you for the comment. The COPD assessment tool has only 8 questions and we wanted to use a more in-depth disease-specific questionnaire. We chose the SF-12 because we didn't want two long questionnaires and we could elaborate more on QoL questions in the interviews.

--Since clofazimine often causes reversible non-melanotic hyperpigmentation that can lead to identifiability and potential stigma (which can vary significantly by ethnogeography), the authors should anticipate stratifying their results by whether clofazimine was part of the treatment regimen or not.

Thank you for this advice, we will look into this in the data analysis, and especially explore it in the qualitative interviews since given the limited sample size, we may not have sufficient power to tease it out in the HRQoL.

Reviewer: 2

1. The sub-section on pages 3-4 on patient reported outcome measures - this seems unnecessarily long and detailed. There is sufficient recognition of the value of these, and the protocol aim could not be achieved in any other way, therefore this extended justification seems an odd inclusion. I suggest that this be amended to instead address HOW understanding patients' QoL on the specific shorter regimen evaluated in PRACTECAL will be useful to implementation, roll-out, and future further iterations of refinements to treatment regiments (i.e., informing development priorities). You can then place greater emphasis on the novelty of an objective measure of QoL impact of the shorter regimen because you have nested this in the trial.

Thank you for your suggestion, we have edited and shortened the section. We have also added a description on how understanding patients QoL will be useful, see page 3, line 91-94.

2. Page 4, lines 20-24 - If the SF36 is the most commonly used, why did you then choose the SF12. Doesn't the reduction in items give less detail for describing QoL? Some justification at this point would be useful.

Thank you for this query, we wanted to use both a general and a disease specific HRQoL instrument, but we also wanted to avoid taking too much of the patients' time so we chose the shorter of the general instruments. We have added a few sentences to address why we chose the SF-12 on page 4, line 140, of the manuscript as follows:

"Being mindful of participant burden we chose to use the SF12 which accurately reproduces the two summary component scores (e.g. physical and mental health) of the SF36 [16]. Additionally, the qualitative interviews will explore further these domains."

3. Page 4, sub-section on qualitative PRO measures in Tuberculosis - I'm not clear about the point of this paragraph. There is a significantly larger literature reporting qualitative data on the general experiences of people with TB (adults and children). There is a smaller, but still robust, literature reporting on qualitative data on the quality of life of patients with TB. If this is a summary of those literatures, then there are notable exclusions. However, I think the authors are still making the more general methodological point that QoL can be studied with qualitative data. If so, see point 1, this just seems unnecessary. It is more important to summarize the literature on qualitative data about patient experiences of MDR-TB treatment, or how MDR-TB treatment has their quality of life (more narrow).

Thank you for the comment. We have addressed this and edited the section, highlighting the paucity of data with this method especially in TB clinical trials and referencing emerging research looking at QoL using qualitative data, page 5, lines 167-170.

4. Page 5, line 8 - The quantitative data are exclusively on QoL impact. To describe this as a study about 'perceptions, expectations and experiences' seems too broad, even though there will also be qualitative data collected. Suggest revising to read 'patient quality of life relevant experiences' instead.

Thank you very much, we have changed the text as suggested, page 5, line 178.

5. Secondary objective 2 - why would we assume that the SGRQ and SF-12 DON'T have utility in TB clinical trials? These are well-established measures. Similarly, of course some qualitative data can be collected as part of clinical trials (this has been done before!), and if done well, this should have utility. Perhaps this is a question of feasibility rather than utility?

But again, still, that seems to be a question that we already have the answer to. Suggest removing this objective.

We recognise that it is rational to expect that the instruments would be useful. However, since we haven't seen TB clinical trials explicitly including these critical aspects especially in non-inferiority designs, we would like to describe them in this trial as one way of advocating for such studies to become routine.

6. Page 5, lines 43-47 - this repeats information already stated in the objectives. Suggest removing here.

This has been removed in the text.

7. Page 5, sub-section on patient and public involvement - this manuscript is reporting on PRACTECAL-PRO, not the overall trial. Were patients and the public involved in PRACTECAL-PRO?

Patient involvement specific to the PRACTECAL-PRO was limited to focus on the cognitive debriefing for the translated questionnaires. As most of these participants are also in the overall trial, and the community advisory boards are more geographic than study-specific, generic engagement around research choices and implementation for the two studies overlap.

8. Page 6, sampling - it is unclear why you would not administer these QoL instruments to all PRACTECAL participants. Why sub-sample at all if you intend to do a descriptive analysis?

There are a few reasons taking into consideration the implications for patients as well as the research staff. Most of the PRACTECAL sites were clinical trial naïve hence we preferred the sites to focus on gaining experience in delivering the main clinical trial before we included the sub-studies. This delayed the start of PRACTECAL-PRO and meant that only some patients could join. Additionally, there are multiple sub-studies running alongside TB-PRACTECAL. We therefore don't consider being part of PRACTECAL automatic consent to join the sub-studies, so expected that some patients may prefer not to join any of the sub-studies while continuing in the main trial. It would have been very difficult and an additional burden to patients and clinic staff to administer the questionnaires to all participants given the resources available.

9. Page 6, healthy controls - I suggest that the point isn't that they are 'healthy', but rather that they experience a QoL that is the average for this population but not people living with TB in this population. I.e., many people would argue that people's average 'healthy' QoL in Scandinavia is higher than these study sites. We're looking for an average person's QoL in this place, not a 'healthy' person's QoL for the comparator, correct?

Thank you for your observation. We are using 'healthy control' as the accepted terminology in QoL research. We are recruiting people who consider themselves generally healthy and accept that it is relative to the context.

10. Page 7, second paragraph under the sub-heading 'Surveys' - this does not seem to be about the instruments but rather a repetition of aim/objectives. Suggest removing.

We have removed this and relocated part of the text to the objectives section, page 5, line 184-187.

11. Page 8, data analysis - suggest describing this as things you will describe, rather than 'objectives include:'

Thank you, we have amended this in the manuscript, page 9, line 334.

12. Page 8, lines 27-33 - repeats from sampling section. Suggest removing.

The repetition on page 9, line 343, has been removed.

13. Page 8, qualitative data - commend the authors for a strong methods description here.

Thank you very much for this comment.

14. Page 8, Discussion section - is this relevant in the absence of findings? It can only duplicate the rationale as outlined in the introduction. Suggest removing.

Thank you for this comment, we have removed any duplication to the rationale. We have kept the text that we feel concludes the main text, page 9, lines 371-377.

VERSION 2 – REVIEW

REVIEWER	Metcalfe, John UCSF
REVIEW RETURNED	12-Jan-2021

GENERAL COMMENTS	The authors have adequately addressed my initial queries. My only remaining suggestion is that healthy controls are endeavored to be matched not only on sex and age, but also socioeconomic status as possible.

REVIEWER	Hoddinott, Graeme Stellenbosch University Faculty of Medicine and Health Sciences, Desmond Tutu TB Centre
REVIEW RETURNED	22-Dec-2020

GENERAL COMMENTS	Overall, congratulations on an important study, and all best wishes for its implementation. I still think that it is a missed opportunity to limit data collection to the SF12 rather than 36 and not include all sites - the world is unlikely to get this opportunity again. Having said that, we all bow to funding/practicality pressures. Minor comments: Page 5, line 26 - "care givers" or "caregivers"? Also, should have a possessive apostrophe: "caregivers' acceptance". Page 8, line 22 - the English for Sotho language is "Sotho", the Sotho for Sotho language is SeSotho. Since the article is English, suggest "Sotho". Else it should also be isiZulu (Zulu for Zulu language). Page 8, line 40 - suggest just "participants' preferred languages" rather than "own native".