PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Using discrete choice experiments to design interventions for heterogeneous preferences: protocol for a pragmatic randomized controlled trial of a preference-informed, heterogeneity-focused, HIV testing offer for high-risk populations
AUTHORS	Ostermann, Jan; Njau, Bernard; Hobbie, Amy; Mtuy, Tara; Masaki, Martha L; Shayo, Aisa; van Zwetselaar, Marco; Masnick, Max; Flaherty, Brian; Brown, Derek S.; Muehlbacher, Axel C.; Thielman, Nathan

VERSION 1 - REVIEW

REVIEWER	Jason Ong Monash University, Australia London School of Hygiene and Tropical Medicine, UK
REVIEW RETURNED	28-Jun-2020

GENERAL COMMENTS	Thank you for the opportunity to review your interesting and important research. Some comments to help improve the clarity of your manuscript. - You mention you want to obtain representative samples - what do you mean by a representative sample and how will you achieve this? Have you got a predetermined quota according to sociodemographic characteristics to fill? - related to the above, could you make it clear (earlier on) what your inclusion/exclusion criteria are? - is there any way to be more objective than self-reports? is photo-verification of test results or test kits feasible? - could you please clarify if the participants that complete the DCE are the same people that join the RCT? This study will be very interesting if you could link their stated preferences with observed behaviours. - I am a little concerned about the assumption that preferences for testing options will not differ between FBW and KMPs. Could you discuss the implications of this assumption? - You have Phases A-E that will be in the same order for all participants - how could you disentangle the ordering effects of these 'interventions' on your final effect? i.e. behaviour in phase E is conditional upon exposure to all preceding phases? how generalizable will your findings be in the 'real world' when not all components of Phases A-E are used?

REVIEWER	Monisha Sharma University of Washington
REVIEW RETURNED	05-Jul-2020

GENERAL COMMENTS	Introduction It would be useful to provide some background on female bar workers and male porters. What are their HIV risks (and prevalence) and their HIV testing coverage? Study population How do the inclusion criteria affect the potential generalizability of the findings, eg. If male porters are a highly mobile population, does restricting to those who will be in the area for 12-15 months impact the selection into the study? Would preferences of bar workers and porters generalize to other high risk groups in Tanzania or other parts of Africa? Reference phase, if participants have already HIV tested in the past 3 months, how likely is it that they would test again in the subsequent intervention phases? It would be useful for the authors to explain the reasoning behind the least preferred option arm Are the participants the same in all phases or are more recruited as the study progresses? What is the rational for keeping the participants the same? If some have already tested in prior phases would this make them less likely to test again? Will participants in the DCE be the same as those enrolled in the trial? Can individuals still test if they do not present the study cards? How will this be handled in the outcome assessment? It would be helpful to have more details on how the study outcomes will be assessed. The authors state “Self report and documentation outcome” will be used. Will self report be higher for certain types of testing, eg HIV self testing, which may be then over reported due to social desirability bias? Do the authors have more details on which HIV testing attributes will be assessed in the DCE? Discussion Some of these sentences can be moved to the methods and described in more detail: “Phone-based follow-up surveys will assess overlap between provider documentation and participant reports of testing and the extent to which participants test outside the study area.” How will this information be reconciled if there is a discrepancy? Which outcome will be used for the main analysis? “Sensitivity analyses will be conducted to describe the effects of

	selection bias on estimates." This is really interesting, how will this be done?

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:
Reviewer: 1
Reviewer Name: Jason Ong

1) You mention you want to obtain representative samples - what do you mean by a representative sample and how will you achieve this? Have you got a predetermined quota according to sociodemographic characteristics to fill?

We thank the reviewer for raising this point. Unfortunately, no prior information on the characteristics of the two populations is available beyond rough estimates of the population size. The sentence introducing the specific recruitment methods now reads: "... the goal is to employ a systematic recruitment approach that minimizes biases." A key study objective is to derive unbiased estimates of testing preferences, that are used to generate preference-informed testing offers for the target population. Thus, the emphasis for our sampling/recruitment approach is on minimizing biases.

2) related to the above, could you make it clear (earlier on) what your inclusion/exclusion criteria are?

We moved the inclusion criteria into a separate section immediately following the description of the two groups of study participants. Eligible study participants are ages 18 or older, reside in Moshi, are able to read, and have no plans to leave the study area during the 12-15 month period following study enrollment.

3) is there any way to be more objective than self-reports? is photo-verification of test results or test kits feasible?

We clarified in the revised protocol paper that test results for the primary study outcome will be documented by counselors after completion of the HIV test. Specifically, for the PRCT, coded invitation cards will be distributed to participants, and HIV tests will be tracked on the basis of cards returned to HIV testing centers in the study area, alongside participants' gender, age, and HIV test results. Self-reports are used to capture tests outside the study area, and tests without cards, including tests during Phases A and B. The primary outcome analysis will use only counselor-documented testing information. (See also comments #8 and #9 by reviewer 2.)

4) could you please clarify if the participants that complete the DCE are the same people that join the RCT? This study will be very interesting if you could link their stated preferences with observed behaviours.

We clarify in the revised version that there may be overlap between the initial DCE participants (who will inform the intervention) and the RCT participants. All PRCT participants will complete the DCE survey to allow for comparisons of stated preferences (DCE survey responses) and revealed preferences (testing decisions).

5) I am a little concerned about the assumption that preferences for testing options will not differ between FBW and KMPs. Could you discuss the implications of this assumption?

We clarify in the revised version of the manuscript that participant type (FBW, KMP) will be included as a covariate in our latent class model. If the distribution of preferences differs significantly between the two types of participants, separate preference-informed testing options may need to be offered. On the other hand, if the preference distributions are broadly similar, the testing offer can be optimized for the joint preferences of both populations. In practice, such an offer would be preferable as it would reduce the dimensionality of the testing offer.

6) You have Phases A-E that will be in the same order for all participants - how could you disentangle the ordering effects of these 'interventions' on your final effect? i.e. behaviour in phase E is conditional upon exposure to all preceding phases? how generalizable will your findings be in the 'real world' when not all components of Phases A-E are used?

The reviewer raises an important question, the answer to which influences estimates of the effect of our preference-informed testing offer on testing uptake, and ultimately, its cost effectiveness.

While we may not be fully able to disentangle the ordering effect, we should be able to estimate whether there is an effect. Specifically, there are three rounds of enrollment into the study, one into Phase A, and two subsequent rounds directly into Phases C and D (see also Figure 1 and notes to Table 2). This approach ensures variation in the exposure to pre-trial intervention components across participants. When modeling testing uptake in each study phase, we will include variables describing participants' exposure to interventions in prior study phases (SMS reminders, testing offers, etc.) as covariates. Significant coefficients on these prior exposures may be indicative of the ordering effects mentioned by the reviewer.

In our study we will not be able to estimate an unconditional effect of Phase E. Ideally, the effect of the incentive would be evaluated in the form of separate study arms. However, we do not consider concurrent incentivized and non-incentivized testing offers viable among potentially closely-knit community members (barworkers in the same bar, porters climbing together). As implemented in this study, the primary purpose of the incentive offer (Phase E) is to observe the testing choices among participants who do not test without an incentive in Phase D.

We clarified these points in the revised version of the protocol paper and added the latter part as a limitation.

Reviewer: 2
Reviewer Name: Monisha Sharma

1) Introduction. It would be useful to provide some background on female bar workers and male porters. What are their HIV risks (and prevalence) and their HIV testing coverage?

We added the following information: "For example, compared to randomly selected community members, FBW and KMP reported 2–3 times as many lifetime sexual partners, higher rates of sexually transmitted illnesses, and higher rates of having sex in exchange for money or gifts, but similar numbers of lifetime HIV tests (median 1–2 across all groups)." In 2003, the HIV prevalence among more than 1,000 FBW in the study area was 19 percent; no HIV prevalence estimates exist for KMP.

2) Study population. How do the inclusion criteria affect the potential generalizability of the findings, eg. If male porters are a highly mobile population, does restricting to those who will be in the area for 12-15 months impact the selection into the study?

It is possible that more mobile participants have different preferences and opportunities for HIV testing. However, we consider this inclusion criterion essential given limited resources for enrollment and follow-up. This is a problem for any longitudinal study of mobile populations. We added this as a potential limitation.

3) Would preferences of bar workers and porters generalize to other high risk groups in Tanzania or other parts of Africa?

We do not believe that the specific preferences of these two study populations will generalize to other high-risk groups in Tanzania or other parts of Africa. However, if this study is successful, it will support the broader use of stated preference methods to systematically elicit the preferences of key target populations, and facilitate corresponding adaptations to HIV testing options. We clarified this point in the discussion section.

4) Reference phase, if participants have already HIV tested in the past 3 months, how likely is it that they would test again in the subsequent intervention phases?

Since 2013, Tanzania's National Comprehensive Guidelines for HIV Testing and Counselling describe specific retesting intervals ranging from 4 weeks to 6 months for most persons at elevated risk of HIV infection. Our own survey of HIV testing sites in the study area revealed that most counselors continue to recommend retesting after 3 months for all clients testing negative for HIV, regardless of risk (unpublished). We added a reference to this in the revised version of the paper.

5) It would be useful for the authors to explain the reasoning behind the least preferred option arm

We clarified in the abstract and in the revised version of the paper that Arm 3 represents an active control arm in which the common option is the best (most preferred, on average) of the testing options given to the participant; the other three options are 'placebo' options that (on average) provide no additional value.

6) Are the participants the same in all phases or are more recruited as the study progresses? What is the rational for keeping the participants the same? If some have already tested in prior phases would this make them less likely to test again?

Please also see comment #6 from reviewer 1, as this point is related to the ordering issue raised by reviewer #1. Ideally, this study could evaluate all tangential intervention components (SMS reminders, invitation cards, incentives) alongside the preference-informed HIV testing offer as part of a multi-arm RCT, however, it would not be possible to enroll a sufficiently large sample into each study arm. Hence the decision was made to let participants progress through incrementally more comprehensive intervention phases. We added this as a limitation.

We show in Figure 1 and highlight in the methods that enrollment will be conducted in three sequential stages (into Phases A, C, and D, respectively). Variation across participants in the stage of their enrollment should allow for separate estimates of the effects of prior intervention components on uptake of testing in the trial. As with exposure to prior intervention components, testing uptake in the respective previous 3 months will be included as a covariate. Given that counselors tend to recommend repeat testing within 3 months (see response to comment #4), it is not clear a priori whether a prior test per se increases or decreases the likelihood of the next test.

7) Will participants in the DCE be the same as those enrolled in the trial?

See comment #1 from Reviewer 1 above. We clarify in the revised version that there may be overlap between the initial DCE participants (who will inform the intervention) and the RCT participants. To allow for a linkage between stated and revealed preferences, all RCT participants will complete the DCE.

8) Can individuals still test if they do not present the study cards? How will this be handled in the outcome assessment?

Yes, participants can test without the study cards. This is the rationale for added self-reports of testing uptake at the end of each study phase. The combination of either counselor-documented or self-reported HIV testing uptake is a secondary study outcome. We clarified this in the revised version of the paper (see "Outcome measure" and also #9)

9) It would be helpful to have more details on how the study outcomes will be assessed. The authors state "Self report and documentation outcome" will be used. Will self report be higher for certain types of testing, eg HIV self testing, which may be then over reported due to social desirability bias?

This is similar to comment #3 by reviewer 1. We clarified in the revised protocol paper that test results for the primary study outcome will be documented by counselors after completion of the HIV test. Specifically, for the PRCT, coded invitation cards will be distributed to participants, and HIV tests will be tracked on the basis of cards returned to HIV testing centers in the study area. Self-reports are used to capture tests outside the study area, and tests without cards, including tests during Phases A and B, and this will be a secondary analysis and help to inform sensitivity analyses around our primary outcome.

10) Do the authors have more details on which HIV testing attributes will be assessed in the DCE?

We clarify in the revised paper that based on prior qualitative work on preference-relevant characteristics of HIV testing options, the DCE will include modifiable HIV testing characteristics in three domains: privacy and confidentiality (e.g. testing venue, different types of counseling), accessibility and value (testing availability, additional services provided), and perceived quality and accuracy (e.g. type of sample for the HIV test).

11) Discussion. Some of these sentences can be moved to the methods and described in more detail:

We now describe the sensitivity analyses in a separate section in the methods

12) "Phone-based follow-up surveys will assess overlap between provider documentation and participant reports of testing and the extent to which participants test outside the study area." How will this information be reconciled if there is a discrepancy? Which outcome will be used for the main analysis?

Please also refer to our response to comment #3 from reviewer 1 and comments #8 and #9 from reviewer 2. The primary analysis will use documented HIV testing as the outcome variable. In our secondary outcome analysis the two measures will be combined using an 'or' condition, i.e., either a documented or a self-reported HIV test will be considered as indicative of an HIV test.

13) "Sensitivity analyses will be conducted to describe the effects of selection bias on estimates." This is really interesting, how will this be done?

We clarify in the revised paper (now in a separate sensitivity analysis section in the methods) that this will be done by modeling attrition as a function of observable characteristics at the time of enrollment; individual-level predictions of the intervention effect (differential predicted probability of testing uptake for each intervention arm relative to the control arm) may then be weighted by the inverse probability of attrition. Differences between the average estimated intervention effect and the attrition-weighted intervention effect characterize the effects of bias from selective retention on our estimates.

VERSION 2 – REVIEW

REVIEWER	Jason Ong Monash University, Australia London School of Hygiene and Tropical Medicine, United Kingdo
REVIEW RETURNED	08-Aug-2020

GENERAL COMMENTS	Thank you for addressing my comments adequately. All the best in carrying out this trial - I will be very interested in the findings from your RCT.