Author’s response to reviews

Title: Strengthening health systems response to violence against women: protocol to test approaches to train health workers in India

Authors:

Sarah R Meyer (smeyer@who.int)
Sangeeta Rege (sangeetatrege@gmail.com)
Prachi Avalaskar (prachi@cehat.org)
Padma Deosthali (padma.deosthali@gmail.com)
Claudia García-Moreno (garciamorenoc@who.int)
Avni Amin (amina@who.int)

Version: 1 Date: 30 Jan 2020

Author’s response to reviews:

Authors’ response to reviewers
Manuscript title:

To the Editors, BMC Pilots and Feasibility Studies

Thank you for the recognition of the contribution of our manuscript, “Strengthening health systems response to violence against women: protocol to test approaches to train health workers in India.” In response to the reviewers’ comments, several changes have been made, as detailed below. We feel that the suggestions have greatly helped improve this manuscript and thank the reviewers for their feedback and comments. The reviewers’ comments are addressed point-by-point in turn below.

Reviewer 1:

1. Implementation science approach. The background section briefly mentions that the evaluation is guided by implementation science; specifically, "the present study will apply implementation science methods to identify aspects of implementation of the Guidelines and related tools, in order to improve understanding of local contextual factors influencing intervention outcomes and support future scale-up" (p.8). It would be helpful if the authors could elaborate on how and which implementation science approaches were used to inform the evaluation design (e.g. a guiding framework?), as this is not readily apparent from the Methods section.
We have elaborated on the theoretical frameworks guiding our research, and added the following text to describe this:

“There are multiple implementation research frameworks; given the research questions addressed in this study, the approach chosen to guide the study design and selection of research methods is a hybrid effectiveness-implementation study, assessing implementation strategy and improving understanding of the contextual factors influencing implementation effectiveness [1]. In addition, we will frame the research with systems-thinking within the context of health-systems strengthening, an analytical approach and conceptual framework that has previously been used to assess positive and negative, intended and unintended consequences, on health systems strength, of a complex health system intervention in Zambia [2, 3].”

2. KAP tool. As the development and validation of the knowledge, attitudes and practice questionnaire is an important part of Objective 1, could the authors provide more information regarding (i) the tool itself (e.g. sample question items, perhaps in the supplementary file), (ii) the results of the pilot phase mentioned on p.14, lines 19-35 (e.g. is there a reference for this?), and (iii) expand on the planned validation methodology (p.15), with corresponding references.

We agree that more information on the KAP survey and validation procedures is warranted.

i) We have now included the full pre-training survey as a Supplementary File, and in addition, have included sample questions in the text of the manuscript.

ii) We have added some examples of changes made through the pilot phase. We have not published results of the pilot phase of research, so we have not added a reference here. The added text is: “For example, the term IPV was removed as it is not utilized in this context, and was replaced with domestic violence [DV], and relevant items were added to reflect local legal frameworks, for example, an item on knowledge of the provisions in the 2006 Protection of Women from Domestic Violence Act. The wording of some items was changed to address issues with comprehensibility that arose in the pilot test – for example, the item “It is demeaning to patients to question them about abuse,” was changed to “It is humiliating to patients to question them about abuse” as demeaning could not easily be translated or understood by respondents in the pilot test.”

iii) We have expanded on the planned validation methodology and added the following text: “In addition, the data will be used to assess various aspects of validity of the instrument, in order to reduce the number of items in the KAP survey and ensure that the scales are validly and reliably measuring relevant constructs. We will assess one aspect of construct validity by assessing convergent validity, assessing correlations of means of scales measuring similar constructs in the survey (i.e. assessing correlation of mean of two different scales for knowledge included in the survey) [4]. We will conduct exploratory factor analysis [EFA] on separate constructs in the survey, to identify i) sub-scales within the larger constructs, i.e. types of knowledge, specific forms of attitudes, and ii) identify redundant items or items that do not correlate well with any factor (determined by having low factor-item loadings). We will assess the number of factors to include using Scree plots and Kaiser criterion, and, after extracting the selected number of factors and conducting the appropriate factor rotation to make the factors more identifiable, we will use the following to guide which items to retain: considering factor loadings (with the general rule-of-thumb of including items with a factor loading of greater than .4), inter-item correlations (with the rule-of-thumb of average correlations of .15 to .5) and item-total
correlations (with the rule-of-thumb of greater than .5) [5]. We will calculate the Cronbach’s alpha of each sub-scale, to explore the extent to which the sub-scales are measuring a single construct; a threshold of an Alpha coefficient of at least .70 is appropriate for the purposes of a validation study [6]. Other aspects of a formal validation study of the scales – for example, criterion validity – are not possible given lack of gold standard measures for knowledge, attitudes and clinical practice in this context.”

3. HMIS confidentiality. In regards to the Health Management Information System, could the authors describe which items will be entered into the data collection form/register and how protection and confidentiality of records will be ensured (e.g. de-identification, encryption, etc.), given the sensitive nature of the information? Per lines 24-27 (p.20), what role will the designated individual from each department have with respect to confidentiality of the register?

We have added the following text to indicate what is collected in the HMIS and how confidentiality is protected: “The items included in the register are: unique patient identification code, age, marital status, presenting signs & symptoms indicating violence, forms of violence woman facing, and details of formal and informal support services provided by healthcare provider. Each department will have a designated person who will ensure the confidentiality of the register. The designated individuals are nurses who are responsible for maintaining the registers, and will keep the register in a locked cupboard, with a protocol for handover during change in shifts.”

4. HSR instrument. As the Health Service Readiness assessment instrument is an important component of Objective 3, could the authors provide (i) sample items from the instrument (perhaps as part of supplementary material), (ii) more detail regarding how and which documents and direct observations will help inform the development of the instrument? Further, is a formal validation planned for this instrument following the pilot phase?

i) We have included sample items from the instrument as a Supplementary File 2.

ii) We have added an example to illustrate documentation and direct observations aspects of the instrument. The following text has been added:

“Based on tools piloted in other settings and understanding of structure of local health facilities, we have proposed source(s) of information for each of the items in the assessment instrument – for example, the item “Is there a written protocol/ standard operating procedures [SOP] for provision of health care to women subjected to domestic and/or sexual violence available in the facility?” will be assessed by asking a facility manager and looking at a copy of the SOP if it exists.”

The items in Supplementary File 2 provide further examples for interested readers.

iii) We will not conduct a formal validation of the instrument during this phase of research. The instrument will be implemented in 18 departments during this phase of research and data analysed to assess feasibility of collecting this level of detail, completeness of data and variability in results. This will inform refinement of items in the instrument, and we will conduct a full validation study for this instrument in the next phase of research.

5. Limitations related to setting. In Limitations, the authors appropriately noted that as the intervention will be implemented in tertiary-care settings, findings may not be generalizable to
primary and secondary-care settings (p.25). Since the patient population of interest may already be hard to reach, it would be helpful to add a sentence acknowledging which populations of women may be missed as a result of implementing the intervention in a tertiary-care setting.

We have clarified this limitation, and do not believe that implementing the intervention in tertiary-level facilities will introduce systematic bias due to the structure of the Indian health care system in these districts. We have added the following text: “However, choice between primary, secondary and tertiary healthcare facilities in these districts is based on proximity and there is unlikely to be systematic bias, such as wealth level or ethnic background, due to sampling at tertiary-level facilities.”

6. Limitations related to study design. In Limitations, it would also be helpful to briefly acknowledge the implications (quant. outcomes) of using an uncontrolled before-and-after design (e.g. historical and maturation bias; parallel or intervening policy or local/contextual changes).

We agree that these are important limitations to highlight, and have therefore added the following text and some relevant references:
“This study is a pre-post study design; therefore, changes in knowledge, attitudes and practices amongst participants of the HCP training, indicated by changes in KAP survey scores, for cannot be fully attributed to the intervention. Without a control group that does not receive training, it is possible that improvements in clinical practice, for example, could be due to other interventions or other contextual changes influencing HCP’s knowledge, attitudes and practices, which are unmeasured in this study design. However, this is a formative research phase, and the pre-post study design was deemed to be the most feasible design, allowing insights into validity of instruments, directions of changes for participants in the training, and feasibility and acceptability of aspects of the interventions.”

Reviewer 2:

1. In the objectives section: "To validate approaches to roll out the training and service delivery improvement activities...". I don't think validate is the right word here. It implies a more quantitative and definitive approach I think. Perhaps say "demonstrate feasibility of..."

We agree that the term “validate” implies a more quantitative approach, and have changed the sentence to read: “To explore feasibility of approaches to roll out….”

2. At the end of the objectives section there is a paragraph about the agencies/organizations involved in the study. This seems odd to put in the introduction/objectives section. Perhaps put it in the methods.

We have moved this section to the Methods section.

3. There is a lot going on in this study. The objectives could each be a separate study on their own. Perhaps a diagram would help readers visualize the study process, timelines, etc.
We have added Figure 1, which includes each study objective and the related methods to clarify the study process for the reader.

4. It is unclear whether this pre-post KAP questionnaire has already been completed/started. Based on the timelines listed at the end of page 13 it seems that it might have already started. In fact, it is unclear exactly how much of this whole study has been complete and how much is still ongoing.

All data collection for the KAP instrument has been completed. We submitted this manuscript to the BMC system nearly a year ago and data collection has been ongoing during that time, as the protocol has approved by all ethics review bodies in 2018. Qualitative data collection is ongoing and data collection for the HSR instrument has not yet started.

5. Page 15 "Exploratory data analysis will be conducted to identify whether baseline levels of attitudes, knowledge and practice vary by sex, role and professional background of respondent, and changes between pre and post-test, and 6-month follow-up will also be explored by relevant socio-demographic variables". More details on the statistical analysis is needed here (at minimum, which statistical test will be used).

We have added the following description to this section: “We agree that more detail on statistical analysis is needed, and have added the following text: “We will compare mean levels of knowledge, attitudes and practices pre, post and post-6 months using paired t-tests or Wilcoxon signed rank t-test (if distribution of paired differences is not normally distributed). We will also compare differences in baseline levels of knowledge, attitudes and practices by different socio-demographic variables (sex, location, role) using t-tests. We will document loss to follow-up, defined as participants who complete the pre/post training survey but not the post 6-month survey. We will analyse if those lost to follow-up are significantly different than those retained in the study on basic socio-demographic variables, to explore potential biases introduced that may impact validity of study findings. Finally, we will use results from the exploratory data analysis to build multivariate models to assess significance of socio-demographic variables (site, sex, age, profession and department) on changes in mean levels of knowledge, attitudes and practices of HCPs.”

6. Page 15, there is a paragraph on psychometrics to be analyzed. There needs to be more detail here. Psychometric measurement studies are complicated and this is the only paragraph explaining the methods of this part of the study. Will this feasibility study be large enough to assess these measures of validity?

We agree that more detail is needed on psychometrics, and have added the following text in response to this and Reviewer 1’s comment on need for more description of psychometric methods: “In addition, the data will be used to assess various aspects of validity of the instrument, in order to reduce the number of items in the KAP survey and ensure that the scales are validly and reliably measuring relevant constructs. We will assess one aspect of construct validity by assessing convergent validity, assessing correlations of means of scales measuring similar constructs in the survey (i.e. assessing correlation of mean of two different scales for knowledge included in the survey) [4]. We will conduct exploratory factor analysis [EFA] on
separate constructs in the survey, to identify i) sub-scales within the larger constructs, i.e. types of knowledge, specific forms of attitudes, and ii) identify redundant items or items that do not correlate well with any factor (determined by having low factor-item loadings). We will assess the number of factors to include using Scree plots and Kaiser criterion, and, after extracting the selected number of factors and conducting the appropriate factor rotation to make the factors more identifiable, we will use the following to guide which items to retain: considering factor loadings (with the general rule-of-thumb of including items with a factor loading of greater than .4), inter-item correlations (with the rule-of-thumb of average correlations of .15 to .5) and item-total correlations (with the rule-of-thumb of greater than .5) [5]. We will calculate the Cronbach’s alpha of each sub-scale, to explore the extent to which the sub-scales are measuring a single construct; a threshold of an Alpha coefficient of at least .70 is appropriate for the purposes of a validation study [6]. Other aspects of a formal validation study of the scales – for example, criterion validity – are not possible given lack of gold standard measures for knowledge, attitudes and clinical practice in this context.”

In terms of whether the feasibility study will be large enough to assess these measures of validity, power calculation for sample size for validation studies are not relevant [7]. There is considerable disagreement regarding adequate sample size for exploratory factor analysis for scale development [8, 9]. Both DeVellis and Netemeyer, Bearden et al. highlight a number of researchers who support sample sizes ranging from 100 to 300 [5, 10]. Using Monte Carlo simulations, Sapnas and Zeller found that a sample size of 100 was adequate to identify a measurement model. We will have a minimum of 170 respondents to the KAP survey, so this is an adequate sample size number of respondents in the study.

7. The sample size is 170 providers. This paragraph needs more details on how this number was reached. Typically pilot studies do not need a formal sample size calculation like a definitive study would, but there should be a clear rationale for the sample size estimation. I think table 1 was meant to explain the sample size but I cannot figure out where the numbers came from.

We have added the following sentence to indicate that the sample size is based on feasibility of how many trainees can be included in the training, rather than a formal sample size calculation, given the purpose and phase of the study:
“Given this is a feasibility study and is not formally testing hypotheses, the sample size is not based on power calculations; instead, we identified the largest feasible sample to include in the training, based on knowledge and understanding of HCP and health manager’s availability to and interest in attending training on response to violence against women.”

8. Page 18: Qualitative assessment plan. There needs to be more information here on the qualitative methods used. For example, maximizing trustworthiness, data analysis methods, sampling methods, and all of the items on the SRQR or COREQ qualitative reporting guidelines.

We have added more description of data processing, ensuring trustworthiness and data analysis. We have added the following text:
“Validity of qualitative data will be approached by through the lens of trustworthiness, including credibility, which is described as ensuring that the process of data collection is “logical, traceable
and clearly documented” [11]. To ensure credibility, for all qualitative data collection activities, IDIs and FGDs will be audio recorded and transcribed by professional transcribers in the local language (Marathi) and/or English. All FGDs will be facilitated by two data collectors, with one focusing on note taking to ensure tape recordings can be cross-checked against hand written notes of the discussion. Local language transcripts will be translated into English and translations cross-checked for accuracy by professional translator. The research team of the PI (i.e. CEHAT) in India will review all transcripts and the research team in WHO, HRP will review select transcripts in English for quality assurance.”

Please see below for our response on sampling methods.

We have reviewed both SQRQ and COREQ and agree that inclusion of all of the items in these checklists is appropriate in the case of reporting the results of a qualitative study. However, in the present manuscript we are presenting a mixed-methods protocol. Therefore, we have added some of the requested items (maximizing trustworthiness, analysis methods) and added some additional items (data processing). Other items in these checklists, including reflexivity, sample description, synthesis and interpretation, will be included in our manuscripts presenting results from the qualitative data collection.

9. Qualitative sample size calculation: I don't think this is how qualitative sample sizes are supposed to work. Usually the process is iterative and the researchers stop sampling once saturation is reached. While purposively sampling for maximum diversity is great, this needs more justification and/or explanation. Also there needs to be a justification for the number of participants per focus group. Some researchers believe that 10 is too many people for a focus group.

We agree that sample size calculation in qualitative research is often iterative, however, the concept of iterative sample size calculation and data saturation is increasingly questioned within discussions of qualitative research design. For example, there are discussions regarding procedures to operationalize the concept of saturation within the actual practices of data collection and thematic analysis [12]. As such, we have added some text to justify our sampling approach: “The sample size calculation is based on feasibility, capacity to capture diverse perspectives, and logistics of implementing data collection requiring busy HCPs to participate in further data collection activities.”

10. Page 18: "we will conduct IDIs with a sample of women who disclose violence during a visit with a health care provider who underwent training, approximately 6 months after the training." This sentence could be revised for clarity. Do you mean the HCPs underwent training or the women who disclosed violence?

We agree and have edited this sentence to read: “we will conduct IDIs with a sample of women who disclose violence during a visit with a trained health care provider.”

11. The authors state that they used the SPIRIT checklist. SPIRIT is a good start but this is not a trial. I think SRQR or COREQ could be included as well.
We have reviewed both the SQRQ and COREQ checklists. Both checklists focus specifically on qualitative research, and have several items related to themes, findings, characteristics of participants, etc, which would require us to fill in the checklist only after all we have collected and analysed all qualitative data. Given this is a protocol manuscript, describing a feasibility study with an intervention, the SPIRIT checklist is the most appropriate checklist.

We appreciate the suggestion of these checklists, and will certainly select one to complete when we report on our qualitative findings.

12. Braun V, Clarke V: To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qualitative Research in Sport, Exercise and Health 2019. Published online