Author’s response to reviews

Title: Discordance in Self-Report and Observation Data on Mistreatment of Women by Providers during Childbirth in Uttar Pradesh, India

Authors:

Arnab Dey (arnab@sambodhi.co.in)
Holly Shakya (hshakya@ucsd.edu)
Dharmendra Chandurkar (dharmendra@sambodhi.co.in)
Sanjiv Kumar (sanjiv.kumar@ihat.in)
Arup Das (arup.das@ihat.in)
John Anthony (john.anthony@ihat.in)
Mrunal Shetye (Mrunal.Shetye@gatesfoundation.org)
Suneeta Krishnan (Suneeta.Krishnan@gatesfoundation.org)
Jay Silverman (jgsilverman@ucsd.edu)
Anita Raj (anitaraj@ucsd.edu)

Version: 1 Date: 19 Oct 2017

Author’s response to reviews:

Response to reviewers

Reviewer 1:

1. Methodology Section

a. Include in the text as to when the observations started and ended? For example, did it start when the woman was in labor and continued until she left the birthing room? Or did it start when she was in the birthing room and ended as soon as baby was born?

Based on feedback from both the reviewers, we have included the duration of observations in the methods section.
b. Clarify the criteria for deciding which provider to observe if there were multiple providers present.

The nurse investigators identified and coded all the health care providers present in the labor room during delivery. For each of the procedures observed during delivery, the code of the provider performing the procedure was marked by the nurse investigators. The nurse investigators then interviewed all the providers identified during the delivery observations using a structured schedule. However, it is worth noting that majority of the deliveries were conducted by one Staff Nurse accompanied by an unskilled birth attendant for support (94%). We have now added this in the methods and results section.

c. In the self-reports, did the women know which provider was being referred to, especially if there were multiple providers?

No. The women were not asked questions specific to providers. However, this does not affect our analysis as we linked provider characteristics to practices based on the procedure and stage of delivery.

d. Did the providers' questionnaire collect information on their reports of their behavior with the women?

We did not capture information providers’ reports of their behavior with women. This is now noted in our limitations.

e. Page 9, line 203. Mistreatment of women is the primary dependent variable and not independent.

Thank you for pointing this out. We have corrected this in the revised manuscript.

f. Page 9, line 207. Clarify what the numbers (0.7) and (22) in brackets are. I think the first is the Cronbach's alpha and the second is a citation. Please make it easy for a reader.

Thanks for pointing this out. We have modified this in the revised manuscript to be clearer for the readers.
g. Page 10, line 216. Explain why 6 items were picked for inclusion for the observation. I suspect this is related to the period of observation.

Of the 17 items in the self-report tool, only six items could be observed during the period of observation. The remaining 11 items of mistreatment could take place before or after the delivery procedure in the labor room, which was beyond the scope of direct observations. We now note this point directly in our description of the measures.

h. It will be useful to readers if the authors discuss the theoretical construct of the mistreatment items. For example, do the 17 items and the 6 items all capture the same underlying construct? One could argue that some are more related to interpersonal behaviors (e.g., abusive language), technical competence (pushing abdomen, pulling baby), and facility inadequacies (e.g., absent providers, discrimination).

Based on this feedback, we have added the underlying theoretical constructs for the 17 items and the 6 items on mistreatment in our description of the measures.

2. Results section

a. Were all the women asked all the 17 items? It is not clear if some of the items pertained to a sub-group of women. For example, the four items—client provided information on treatment, information on problem, advice on avoiding illness, consent before treatment—clarify if they refer to some specific treatment or condition that pertain to some women. A substantial proportion of women have self-reported mistreatment on these elements.

Yes. All women were asked all the 17 items. We have clarified this while describing the measures. However, we agree that the term treatment and problem are a bit misleading. The measures were asked to all women and enquired whether the provider told them about the procedures and took consent before performing procedures. We have revised the same in table 1.

b. The authors should comment on the reasons why Cronbach's alpha varied so much between the three modalities (0.72 on the 17 point self reports; 0.64 on the 6 item self report; and .47 on the 6 item observation). Given that the alpha on the observation was less than 0.5, how well do the authors think that the 6 items on the observations modality captures mistreatment? Relatedly was the reduced alpha due to the number of items? Should more items have been included?
We have added text in the discussion explaining the alpha values for the three measures as follows:

The three measures of mistreatment studied in this paper viz. 17 item self-report, 6 item self-report and 6 item observation had Cronbach alpha’s of 0.72, 0.64 and 0.47 respectively. While the self-reported measures show good internal reliability, the reliability of the observed mistreatment is relatively low. Future research on mistreatment can include more items to the observations to test the internal reliability of observed mistreatment.

c. Page 13, lines 293-303. The authors should comment on if the observations of poor provider behavior were due to poor skills of the provider; and second, if the observer’s knowledge of the SBA status of the provider might have resulted in those behaviors being noted more.

We describe the associations between provider characteristics and observed mistreatment in the next paragraph. Observed mistreatment by providers was found to be associated with age of providers and SBA training received by the providers.

The point on observer’s knowledge on SBA is also very crucial and we describe in the discussion section that observers being primed to look for poor quality of care might have resulted in those behaviors being noted more.

d. Was there any link with pregnancy complications that might have had an effect on the self-reports? Especially since health complications of mother and baby seem to be correlated with reports of mistreatment.

We included pregnancy complications in the models for self-reported mistreatments but did not find any association between self-reported mistreatment and complications during pregnancy.

3. Discussion Section

a. Situate findings from the study with those reported elsewhere; for example, the 2015 Abuya et al study in Kenya which also analysed data from observations and self-reports of mistreatment.

Thanks for this suggestion. We have included text in the discussion section comparing findings from our study with the Abuya et. al. work in Kenya.
b. It will be helpful for the authors to comment/discuss the rationale or benefits of collecting data through self reports and observations. In different parts of the paper, the authors allude to the utility of different data sources. They could draw upon the literature on quality of care of FP where relevant. What is the authors' opinion or preference? Do they recommend self report as an explanation of service utilization? Observations to assess provider adherence for service protocols?

While we have some thoughts on this, we do not feel the findings of the current study provide sufficient evidence to justify what issues are more validly captured based on observations and what issues are more validly captured via self-report. This is an important area for future research to consider. For now, we recommend inclusion of both types of measurement, observation and patient self-report, and perhaps provider reports as well, which we were not able to include in the current study. We note these points in our conclusion and limitations.

c. Could the authors comment on the types of health complications that mothers and babies had? This will tie in with the lines 94 and 95 in the background section.

Thanks for this feedback. We have described maternal and neonatal problems while describing our measures.

d. The authors should comment on to what extent do women's self-reports might encapsulate the entire birth experience and not be limited to specific provider behaviors despite asking them to recall specific behaviors? This will be especially true for those women who had poor health outcomes.

We recognize that there is a chance of women encapsulating the entire birth experience while reporting mistreatment. While some of the items are more specific like beaten / slapped or threatened during delivery, there is a chance that others items may be influenced by poor health outcomes for the mother or the baby. We discuss this in the limitations section of the paper.

e. I was glad to see that the authors acknowledge that causality cannot be determined.

f. It will be great if the authors identify potential interventions to reduce mistreatment. They suggest that trained providers behave better than others; could they speculate as to what type of training--it seems to me that training in interpersonal skills will be just as important as technical skills especially as what the women report have tended to be the emotional aspects of the care giving process. Similarly, can interventions be directed towards women, families
and communities such that they know what good quality maternity care is; and that therefore they can expect it and demand good quality care.

Thanks for this suggestion. We have included a paragraph on potential interventions in the conclusion section.

Reviewer 2:

1. The items included in the observation are very limited—mainly verbal and physical abuse which tend to have lower prevalence than other domains of mistreatment. It will have been useful to see discordance/concordance on the items that were dropped such as supportive care from provider and lack of privacy during delivery. In addition, three of the items among the 6 items that were both observed and self-reported likely have different prevalence from self-reports and observations because the women and observers had different vantage points. It is possible that a woman might not notice during second stage if the provider pushed her abdomen forcefully or if her baby is forcefully pulled out (depending on the force) while this may be obvious to the observer. Even then, what is considered forcefully might be very subjective even among observers. Similarly, if the woman experienced a problem due to unavailability of the provider, the observer might miss it if it is not an obvious problem. It is therefore not surprising these have the lowest Kappa, with the former being more likely to be observed and the latter more likely to be reported. Thus, while under-reporting is a potential reason for the discordance, the other reason is women, as well as providers in the last case, may actually not be identifying or interpreting the issue as something inappropriate. The limitations of these measures and their potential effects of the results should therefore be discussed.

We acknowledge the issue of different vantage points for women and observers which is especially relevant in the mistreatment items highlighted by the reviewer. We have included text in the discussion to explain the discordance of these items and also discussed the limitation in methods to different vantage points.

2. Why were only these six mistreatment items observed?

The direct observation only included those items that could be observed during the period of observation i.e. point a woman was admitted to the labor room and continued until she left the labor room. All other items of mistreatment which could have occurred before or after the period of observation were not included in the direct observation. We now note this in the measures section of our paper.
3. It looks like the additional items in the version of the mistreatment tool used in this study included more communication related items than other domains, and communicating information to women seems very poor. Might this account for the much higher rate of mistreatment based on the longer tool? In the discussion of the difference in prevalence of mistreatment based on the different tools, it is important to highlight the major difference in the content of the tools to help identify what is contributing to the big difference in the prevalence of mistreatment.

We acknowledge that some of the additional items included in the current study had very high incident rates and was one of the main reasons for substantial difference in the prevalence of mistreatment between the two studies. We have added text in the discussion to explain this.

4. How were women recruited? Were they recruited before they went on admission or at the time of admission? Were all women on the ward during the observation period observed or only some were observed? Did you obtain consent of both providers and mothers for the observation?

All women admitted to the labor room for delivery during the period the team stayed at the facility were recruited for observations. Written consent was taken both from the women as well as from the providers. We have detailed this in the methods section.

5. What was the period of observation? Did it end after a woman delivered or continued till after discharge from the facility. This is important to assess if some of the predictors should be included, and to tease out the role of temporal sequence in the discussion on causality. For example if observations ended at delivery then postpartum and neonatal complications as well as length of stay at the facility post-delivery occurred after the mistreatment, which will be consistent with your previous work, and should not be considered predictors as shown in the logistic regression in tables 4 and 5.

We have indicated the period of observation in the methods. However, measures on postpartum and neonatal complications and length of stay at the facility post-delivery were captured through self-reports from the follow-up survey.

While postpartum and neonatal complication temporally may have manifested after the mistreatment, those complications are causally part of the birth experience as a whole and not actually a discrete event that occurred after the mistreatment. So in a temporal sense we are considering the birth event to include delivery and postpartum outcomes.
6. Do you have information on sex of providers or all were females? And also on number of providers present during the delivery, day of the week, time of observations, and number of laboring women on the ward at the time of the observation. It will be interesting to see if these are associated with mistreatment.

98% of the deliveries were conducted by female Staff Nurses or Auxiliary Nurse Midwives (ANMs). 94% of the deliveries were conducted by 1 staff nurse or ANM supported by an unskilled birth attendant. We have mentioned these points in the results section.

However, we do not have information on the day of week, time of observation, or the number of laboring women in the ward at the time of observation. These are really interesting suggestions and we can plan to add these in the next round of assessments.

7. Table 3 is not very clear. Usually the table is set up to show group differences for a particular outcome; eg % reporting mistreatment or observed to be mistreatment. The percentages seem to be set up this way, which makes the second column redundant, since we can infer one from the other: % not reporting mistreatment = 100% minus % reporting mistreatment. You however note that "p-values assess differences between groups who experienced and did not experience mistreatment during childbirth on the given variable, based on chi-square analyses for categorical variables and t-tests for continuous variables." This seems true for the continuous variables (I am assuming 26.8 and 26.7 are the mean ages for those who reported mistreatment and no mistreatment). But it is confusing for the categorical values. For example, does the 79.5% for literacy represent the proportion of literate women who were mistreated or the proportion of literate women among those who were mistreated? I assume the later, with the chi-square assessing the differences in mistreatment between literate and illiterate women. That is you are testing for the difference between 79.5 and 75.2 and not between 79.5 and 20.4. This interpretation seems consistent with the description of the results in the text. It took me a while to understand what you were comparing in table and I think others might have the same problem. So please check this to be sure what you are comparing and set up the table to make it easy for the reader to understand. The continuous variables in this table might be better recoded into categorical variables for consistency of the comparison. This might also help to examine the associations better, given the effects of age and age at marriage may not be linear. The table can be also be made less congested by excluding the columns on % not reporting mistreatment, since the % in the two columns add up to 100.

Based on this comment, we have re-structured table 3 and removed the column reporting % of women not reporting mistreatment. We have however, retained mean and standard deviations of the continuous variables in the table to maintain comparability between the independent variables used in the regression models in tables 4, 5 and 6.
8. References 14 and 22 seem to refer to the same publication; please check

Thanks for pointing this out. We have modified this in the text as well as in the references.