

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Effect of a multifaceted intervention on the utilization of primary health care for maternal and child health care in rural Nigeria: A quasi-experimental study
<b>AUTHORS</b>	Okonofua , F; Ntoimo, Lorretta; Yaya, Sanni; Igboin, Brian; Solanke, Ojuolape; Ekwo, Chioma; Johnson, Ermel; Sombie, Issiaka; Imongan, Wilson

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Eboreime, Ejemai University of Alberta Faculty of Medicine and Dentistry, Department of Psychiatry
<b>REVIEW RETURNED</b>	06-Mar-2021

<b>GENERAL COMMENTS</b>	<p>Eboreime, Ejemai University of Alberta Faculty of Medicine and Dentistry, Department of Psychiatry</p> <p>This is an interesting article reporting a complex intervention designed to increase the access of rural women to antenatal, intrapartum, postpartum, and childhood immunization services offered in primary health care facilities.</p> <p>I have a few comments which should make this better for readership.</p> <p>Study settings: The information provided under study setting is sparse, some may not be relevant. What may be more informative is the structure of the health system, number and categories of health facilities, referral linkages, etc. The population of the LGAs may be useful as well, in addition to information about health workers at the PHC level. What are WDCs? How are they structured? How do they function?</p> <p>Methods</p>
-------------------------	--

<b>REVIEWER</b>	Goldblatt, Peter University College London, UCL Institute of Health equity
<b>REVIEW RETURNED</b>	05-Apr-2021

<b>GENERAL COMMENTS</b>	This paper gives a good account of an intervention to persuade more women in rural areas to give birth in primary health centres by making multi-faceted improvements in the offer provided by these centres. The design of the evaluation, using a baseline and
-------------------------	--

	<p>an end line survey is appropriate. The analysis of the principal benefits of the intervention, using odds ratios (Table 3) is sound.</p> <p>However, I found many aspects of the statistical tables hard to follow and clarification of the messages in these tables is necessary, as follows:</p> <p>Table 1</p> <p>It is unclear what the purpose or interpretation of column 1 (All) is, since this is the average of the independent baseline and end line surveys. The purpose of the intervention is to compare the results of the surveys, not take an average over the two surveys.</p> <p>In two rows the figures in parentheses are labelled as standard deviations, but not in all the other rows. It would be better to have a footnote to the table saying what is parentheses in each row.</p> <p>The sub-table on religion is slightly odd - over half those responding are in the "other Christian" category. Some explanation is needed of why it was not possible to disaggregate this number, given that the much smaller group of "Catholics" were separately identified.</p> <p>Table 2</p> <p>It is unclear from the text and footnotes why the numbers in each sub-table sum to different totals and why all these totals differ from the numbers in the baseline survey (1,408) and end line surveys (1,411) on which they should be based.</p> <p>For postnatal care and immunization, it is not clear to which babies these refer - all other sub-tables are indicated to relate to the last birth.</p> <p>Table 3</p> <p>The heading says that the table shows odds - but the text says OR (presumably, odds ratios). This mismatch needs to be clarified.</p> <p>In most sub-tables, one category has "RC" in parentheses. This is presumably the "reference category" for odds ratios. This needs to be indicated in a footnote to the table. For age, ages at marriage and number of children no reference categories are defined. These should be indicated in a footnote.</p> <p>Table 4</p> <p>The heading in this table refers to "most recent birth". Some explanation is needed of whether or how this differs from "last birth" in Table 3.</p> <p>It is unclear how the numbers of respondents in this table have increased by orders of magnitude from previous tables- 2294 in the baseline survey, compared to 1408 in Table 1 and 3698 in the end line survey compared to 1411 in Table 1.</p> <p>It is unhelpful in evaluating the benefit of an intervention to record the change between baseline and end line as "difference of baseline minus end line". This means that all increases are shown as negative numbers and vice versa. It is also inconsistent with differences in Table 2.</p>
--	---

	<p>Table 5</p> <p>It is unclear where the sample numbers (532 and 81) for each survey come from - they do not relate to the numbers of non-users of primary health centres shown in Table 2. Some explanation is needed.</p> <p>Given the small numbers of respondents in the end line survey, it is not clear that any of the differences are statistically significant. It would be better to show a statistic based on a chi-squared test rather than a change in percentage points. As in Table 4, positive increases in percentages are unhelpfully shown as negative numbers.</p> <p>Other comments</p> <p>There are a few obvious typos in the paper e.g.</p> <p>Page 3, line 49</p> <p>"pregnant women use in rural Edo, used" should be "pregnant women in rural Edo used"</p> <p>Page 5, line 22</p> <p>"obtain treatment at PHCs without delay from an inability to pay" should probably be "obtain treatment at PHCs without being deterred by an inability to pay" or something similar.</p> <p>There are also some instances where greater clarity is needed e.g.</p> <p>Page 4 lines 46 and 49</p> <p>After "from 3,462 households" and "from 3,116 households" add "in these 20 communities"</p> <p>Page 7, line 44</p> <p>Explain what alpha means in "Alpha was set at 0.05" - for example "95 percent confidence intervals were calculated".</p> <p>.</p> <p>se</p>
--	--

## VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Ejemai Eboreime, University of Alberta Faculty of Medicine and Dentistry

Comments to the Author:

This is an interesting article reporting a complex intervention designed to increase the access of rural women to antenatal, intrapartum, postpartum, and childhood immunization services offered in primary health care facilities.

I have a few comments which should make this better for readership.

Study settings:

The information provided under study setting is sparse, some may not be relevant.

What may be more informative is the structure of the health system, number and categories of health facilities, referral linkages, etc. The population of the LGAs may be useful as well, in addition to information about health workers at the PHC level.

Response: Some information on the structure of Nigeria's health system were provided in the introduction (line 32-46). We have now provided additional information in the setting. Thanks.

What are WDCs? How are they structured? How do they function?

Response: This has been explained.

Reviewer: 2

Prof. Peter Goldblatt, University College London

Comments to the Author:

This paper gives a good account of an intervention to persuade more women in rural areas to give birth in primary health centres by making multi-faceted improvements in the offer provided by these centres. The design of the evaluation, using a baseline and an end line survey is appropriate. The analysis of the principal benefits of the intervention, using odds ratios (Table 3) is sound.

However, I found many aspects of the statistical tables hard to follow and clarification of the messages in these tables is necessary, as follows:

Table 1

It is unclear what the purpose or interpretation of column 1 (All) is, since this is the average of the independent baseline and end line surveys. The purpose of the intervention is to compare the results of the surveys, not take an average over the two surveys.

Response: Column 1 refers to all respondents (baseline and endline). However, we have deleted it to concentrate on comparing the two surveys. Thanks.

In two rows the figures in parentheses are labelled as standard deviations, but not in all the other rows. It would be better to have a footnote to the table saying what is parentheses in each row.

Response: We appreciate this observation. Notes have been added to explain what the figures are.

The sub-table on religion is slightly odd - over half those responding are in the "other Christian" category. Some explanation is needed of why it was not possible to disaggregate this number, given that the much smaller group of "Catholics" were separately identified.

Response: The category of other Christian was not disaggregated because there many other Christian denominations in the LGAs, and not a single one is as large as Catholics. Thus, we did not collect data separately for each of the other Christian denominations.

Table 2

It is unclear from the text and footnotes why the numbers in each sub-table sum to different totals and why all these totals differ from the numbers in the baseline survey (1,408) and end line surveys (1,411) on which they should be based.

Response: The difference comes from non-response. We have provided more explanation on this.

For postnatal care and immunization, it is not clear to which babies these refer - all other sub-tables are indicated to relate to the last birth.

Response: Thanks. They all refer to the last birth. We have included that.

### Table 3

The heading says that the table shows odds - but the text says OR (presumably, odds ratios). This mismatch needs to be clarified.

Response: This has been addressed.

In most sub-tables, one category has "RC" in parentheses. This is presumably the "reference category" for odds ratios. This needs to be indicated in a footnote to the table. For age, ages at marriage and number of children no reference categories are defined. These should be indicated in a footnote.

Response: This has been done.

### Table 4

The heading in this table refers to "most recent birth". Some explanation is needed of whether or how this differs from "last birth" in Table 3.

Response: They refer to the same birth. It has been corrected to most recent birth in the entire manuscript for consistency. Thanks.

It is unclear how the numbers of respondents in this table have increased by orders of magnitude from previous tables- 2294 in the baseline survey, compared to 1408 in Table 1 and 3698 in the end line survey compared to 1411 in Table 1.

Response: The numbers in Table 4 and 5 do not refer to the number of respondents. They refer to the number of responses. This was stated in the analytical approach section. The questions that produced these results had multiple option responses. A respondent is permitted to select multiple answers.

It is unhelpful in evaluating the benefit of an intervention to record the change between baseline and end line as "difference of baseline minus end line". This means that all increases are shown as negative numbers and vice versa. It is also inconsistent with differences in Table 2.

Response: Thanks. These are percentage point difference which do not necessarily show the impact of the intervention. However, we have deleted them from Tables 2, 4, and 5.

### Table 5

It is unclear where the sample numbers (532 and 81) for each survey come from - they do not relate to the numbers of non-users of primary health centres shown in Table 2. Some explanation is needed.

Response: These are not sample numbers. They are number of responses to the multiple choice questions. See response to comment on Table 4 above.

Given the small numbers of respondents in the end line survey, it is not clear that any of the differences are statistically significant. It would be better to show a statistic based on a chi-squared test rather than a change in percentage points. As in Table 4, positive increases in percentages are unhelpfully shown as negative numbers.

Response: We have conducted chi-square test and fisher exact test where the value in a cell is less than 5. Thanks.

Other comments

There are a few obvious typos in the paper e.g.

Page 3, line 49

"pregnant women use in rural Edo, used" should be "pregnant women in rural Edo used"

Response: Thanks. This has been corrected.

Page 5, line 22

"obtain treatment at PHCs without delay from an inability to pay" should probably be "obtain treatment at PHCs without being deterred by an inability to pay" or something similar.  
Response: Thanks. This has been edited.

There are also some instances where greater clarity is needed e.g.

Page 4 lines 46 and 49

After "from 3,462 households" and "from 3,116 households" add "in these 20 communities"  
Response: Thanks. This has been edited.

Page 7, line 44

Explain what alpha means in "Alpha was set at 0.05" - for example "95 percent confidence intervals were calculated".

Response: This has been explained.

## VERSION 2 – REVIEW

<b>REVIEWER</b>	Goldblatt, Peter University College London, UCL Institute of Health equity
<b>REVIEW RETURNED</b>	17-Aug-2021
<b>GENERAL COMMENTS</b>	<p>The authors have addressed the majority of comments satisfactorily. However, the issues raised around Tables 4 and 5 are not fully addressed.</p> <p>1) it is not clear the number of women who were asked questions about reasons for "using or not using a PHC for skilled care" (see last two sentences of "variables and measures" on page 8. On the face of it, Table 4 entitled " Reasons for use of PHC for delivery care (most recent birth)" should be based on those whose place of delivery was "PHC in the project community" shown in Table 2. However, the figures for this in Table 2 are 417 at baseline and 802 at endline, whereas in Table 4 "Good quality service" received</p>

	<p>451 and 882 positive responses at baseline and endline, respectively i.e. more positive responses than attendees.</p> <p>By contrast, in Table 5 relating to non-use of PHC for delivery, the sum of all positive responses to questions is substantially less than the total number of non-users in Table 2 i.e. most people did not give a reason for non-use. This is not commented on in the paper, in terms of limitations of the survey design.</p> <p>2) The paper now clearly states that the sample numbers in Tables 4 and 5 are the number of responses, not the number of respondents. However, this is not quite accurate. These numbers (for example 2294 for the baseline in Table 4) are the column sums i.e. the total number of positive responses to all the questions in Table 4. It does not include negative responses - the number of which can only be known if non-respondents are excluded.</p> <p>3) The percentages shown in Tables 4 and 5 are based on the column totals i.e. using the sum of positive responses as the numerator. In this way they show within column distribution of positive responses rather than the percentage of people giving a positive response to each question. Statistically, the percentages in each column are therefore correlated as single items in a multinomial distribution.</p> <p>4) Based on this, it is not clear from the paper how independent p values have been calculated in the final columns of Tables 4 and 5. A clear explanation of this is required (possibly in a short statistical appendix).</p>
--	---

## VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

1) it is not clear the number of women who were asked questions about reasons for "using or not using a PHC for skilled care" (see last two sentences of "variables and measures" on page 8. On the face of it, Table 4 entitled " Reasons for use of PHC for delivery care (most recent birth)" should be based on those whose place of delivery was "PHC in the project community" shown in Table 2. However, the figures for this in Table 2 are 417 at baseline and 802 at endline, whereas in Table 4 "Good quality service" received 451 and 882 positive responses at baseline and endline, respectively i.e. more positive responses than attendees.

Response: The women who responded that they used a PHC responded to the question on why they used it, whereas those who did not use a PHC were asked why they did not use a PHC. The observation about the number of positive responses exceeding the number of attendees who used a PHC in the project community has been noted and revised. Please, see more explanation in the response to comment #2. Thanks.

By contrast, in Table 5 relating to non-use of PHC for delivery, the sum of all positive responses to questions is substantially less than the total number of non-users in Table 2 i.e. most people did not give a reason for non-use. This is not commented on in the paper, in terms of limitations of the survey design.

Response: Thanks. This has now been acknowledged as a limitation.

2) The paper now clearly states that the sample numbers in Tables 4 and 5 are the number of responses, not the number of respondents. However, this is not quite accurate. These numbers (for example 2294 for the baseline in Table 4) are the column sums i.e. the total number of positive responses to all the questions in Table 4. It does not include negative responses - the number of which can only be known if non-respondents are excluded.

Response: We have presented the number of positive responses to each reason and the percentage for each reason is now calculated over the number of respondents who used or did not use a PHC in the project community instead of the sum of the positive responses. The number of responses changed from what was originally reported because we observed the error in our initial calculation for the reasons for use and non-use. This has been corrected and all the analyses for reasons for use and non-use are limited to only those who used or did not use the PHCs in the project communities

3) The percentages shown in Tables 4 and 5 are based on the column totals i.e. using the sum of positive responses as the numerator. In this way they show within column distribution of positive responses rather than the percentage of people giving a positive response to each question. Statistically, the percentages in each column are therefore correlated as single items in a multinomial distribution.

Response: We have revised this to show the number of positive responses to each reason, and the corresponding percentage of the total number of women who used or did not use a PHC at both baseline and end line.

4) Based on this, it is not clear from the paper how independent p values have been calculated in the final columns of Tables 4 and 5. A clear explanation of this is required (possibly in a short statistical appendix).

Response: Each reason for use or non-use was generated as a dummy variable with yes as the positive response and no otherwise. The p-values were derived from a cross-tabulation of each reason by the survey period (baseline and endline). The results tables obtained from the analysis are



now included as supplementary material. More explanation about this has been included in the analytical approach in the methods section.

### VERSION 3 – REVIEW

<b>REVIEWER</b>	Goldblatt, Peter University College London, UCL Institute of Health equity
<b>REVIEW RETURNED</b>	22-Nov-2021
<b>GENERAL COMMENTS</b>	<p>The authors have made the substantial changes to Tables 4 and 5 requested in the previous review and have amended the text to adequately reflect any potential non-response bias.</p> <p>They have also made general, minor improvements to the English to improve the acceptability of the paper for publication.</p>