PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Effect of continuum-of-care intervention package on improving contacts and quality of maternal and newborn health care in Ghana: a cluster randomized controlled trial
AUTHORS	Okawa, Sumiyo; Gyapong, Margaret; Leslie, Hannah; Shibanuma, Akira; Kikuchi, Kimiyo; Yeji, Francis; Tawiah, Charlotte; Addei, Sheila; Nanishi, Keiko; Oduro, Abraham; Owusu-Agyei, Seth; Ansah, Evelyn; Asare, Gloria; Yasuoka, Junko; Hodgson, Abraham; Jimba, Masamine

VERSION 1 – REVIEW

REVIEWER	Debra Singh Kimanya-Ngeyo Foundation for Science and Education, Uganda
REVIEW RETURNED	10-Aug-2018

GENERAL COMMENTS	The paper is generally interesting and adds new knowledge. My only issues are 1. The results are difficult to read and interpret - it would be good if they could present them in a clearer manner. "the effect of the intervention on having adequate contacts with high quality care. After the intervention, 12.6% of women in the intervention group had adequate contacts with high-quality care during ANC, with no significant effect in the intention-to-treat" - many of these kinds of statements made it difficult to read. So if they could rethink how to make the information accessible and more readable that would help 2. The other issue is that they introduce the issue of insurance at the end -stating that national health insurance would be the solution to this issue. I would have liked to understand more about the current health insurance scheme and what it offers in the perinatal period. Also were there factors about those that take out insurance that make them more likely to do well generally.

REVIEWER	Onaedo Ilozumba Vrije Universiteit Netherlands
REVIEW RETURNED	19-Nov-2018

GENERAL COMMENTS	A well-written manuscript that considers the important issue of not only increased contact with health care workers but the quality of those contacts. My main comment is that the authors include study population includes a variation of ages and women with limited or no formal

	education. It would be interesting to see if there are differences across ages and education, particularly within the intervention group. My understanding is that the DiD analyses only adjusted for access to district hospitals as a cofounder and not these two factors. Although the protocol extensively discusses study site selection it would be useful to provide a brief justification/explanation of the study site selection within this manuscript. What was the range of contacts with health care professionals during the ANC period? The main outcome measure of "at least 4" indicated there might have been more contacts for some women. However, this is not shown in any of the tables. In addition, this study was conducted prior to the updated WHO recommendations of 8 ANC visits. However, it would be interesting to see the influence of this on the studies results or read the authors reflections on 8 visits at ANC as opposed to at least 4. The Ghana National Safe Motherhood Service protocol is from 2008 so naturally does include such updates.

REVIEWER	Khitam Muhsen Tel Aviv University, Israel
REVIEW RETURNED	02-Feb-2019

GENERAL COMMENTS	This is an interesting and important study. Specific comments Please add to the abstract some details on the main statistical analysis. Please provide point estimates of the study outcomes at baseline and post-intervention in the abstract. Please rephrase the study aim according to PICO (population, intervention, comparison and outcome). Methods Please add more details on the control group. Did the control group receive the standard of care (what did it include?) or other intervention? It is not clear if the assumptions of the difference-in difference (DiD) method were met. Some characteristics of the intervention and control groups differed between the baseline and follow-up assessments, including parity, wealth quintiles and health insurance. It is worth assessing the impact of these factors, especially having health insurance on the study outcomes. Page 18, table 5: why this analysis was limited to the intervention group? Are the determinants of adequate contacts with high quality care different in the control group? An important message herein is reducing social disparities and increasing health insurance coverage to reduce gaps in adequate contacts with high quality care in the treatment of pregnant women and their babies.

REVIEWER	Lifeng Lin Florida State University, USA

REVIEW RETURNED	09-Feb-2019

GENERAL COMMENTS	This manuscript presents results from a cluster randomized controlled trial conducted in Ghana on the effect of the CoC intervention. I have focused on the statistical analyses. The analyses were generally performed well. My comments are as follows. First, on page 7, the authors created 16 pairs from the 32 clusters, accounting for the population, the volume of delivery, and the number of midwife. I'm wondering if the pairing was within each of the 3 sites (Navrongo, Dodowa, and Kintampo). Also, were the clusters randomly paired? If yes, how was the randomization performed while the population, the volume of delivery, and the number of midwife needed to be similar for each pair of clusters? The authors may present a little more details about the pairing, e.g., providing cluster-level characteristics. Second, on page 12, the authors performed the logistic regression with several independent variables, including study site, parity etc. Were all independent variables treated as categorical variables? Was there any continuous variable? If all independent variables were categorical, the authors may consider presenting the detailed categories classified for each variable. Third, Table 2 on page 13 shows that religion and wealth quintiles were significantly different in the control and intervention groups, which may confound the final results. Although the authors mentioned that the intervention group had more Muslim and wealthy women than the control group, I think they may discuss more on these significant differences. What were the potential impact of these two variables on the effect of the CoC intervention? As wealthier women tended to receive better healthcare services, was wealth quintile possible a confounder? Also, was religion spatially related to the clusters' locations? These may be discussed as limitations in this last section. Minor comments: On page 11, the second line from the bottom. Should "efficacious" be "efficacy"? On page 12, I think it is better to put the sentences at the top (for the per-protocol analysis, … excluded 238 women … and 134 women …) in the results section. On page 12, the authors mentioned that they used "robust variance estimate" to control for potential correlations within clusters. I am not very clear what was the exact robust estimate used for variance and how it was related to the potential correlations. The authors may provide some citations on this or give a bit more details.

VERSION 1 – AUTHOR RESPONSE

Response to Reviewers' Comments

Reviewer: 1

Reviewer Name: Debra Singh

Institution and Country: Kimanya-Ngeyo Foundation for Science and Education, Uganda

1. The results are difficult to read and interpret - it would be good if they could present them in a clearer manner. "the effect of the intervention on having adequate contacts with high quality care. After the intervention, 12.6% of women in the intervention group had adequate contacts with high-quality care during ANC, with no significant effect in the intention-to-treat" - many of these kinds of statements made it difficult to read. So if they could rethink how to make the information accessible and more readable that would help
➢ **Response to reviewer** We explained the definitions of intention-to-treat and per-protocol designs more clearly in the statistical analysis section. (L243-253), and revised the results of these analyses. (L306-318)

2. The other issue is that they introduce the issue of insurance at the end -stating that national health insurance would be the solution to this issue. I would have liked to understand more about the current health insurance scheme and what it offers in the perinatal period. Also were there factors about those that take out insurance that make them more likely to do well generally.
➢ **Response to reviewer** We explained the current national health insurance scheme in detail, and added existing evidence from two of our previous studies that women with health insurance membership were more likely to have four ANC visits and delivered at health care facilities. (L102-106).

Reviewer: 2

Reviewer Name: Onaedo Ilozumba

Institution and Country: Vrije Universiteit, Netherlands

1. My main comment is that the authors include study population includes a variation of ages and women with limited or no formal education. It would be interesting to see if there are differences across ages and education, particularly within the intervention group. My understanding is that the DiD analyses only adjusted for access to district hospitals as a cofounder and not these two factors.
➢ **Response to reviewer**

In the revised analysis, we adjusted the following variables as potential confounders: study site, living in a sub-district with a district hospital, age, education, marital status, parity, wealth status, religion, and the status of the national health insurance membership. (L213-219, L238-239, L257-260)

2. Although the protocol extensively discusses study site selection it would be useful to provide a brief justification/explanation of the study site selection within this manuscript.

- **Response to reviewer**

We selected the three study sites which were rural area, and had diverse socio-economic and ecological background and healthcare system challenges, so that we could enhance the generalizability of the study findings in rural settings of Ghana. In addition, the study sites had Health Research Centers under Ghana Health Service, and these centers operated the Health and Demographic Surveillance System. We considered that such research infrastructure could be beneficial for quality control of intervention and surveys. (L125-131)

3. What was the range of contacts with health care professionals during the ANC period? The main outcome measure of "at least 4" indicated there might have been more contacts for some women. However, this is not shown in any of the tables.

- **Response to reviewer**

In our study sample, the median ANC contacts of women with healthcare providers was 5 contacts (interquartile range 4 – 7). We used a cutoff at four contacts because it follows the national guidelines, and was the standard monitoring indicator at the time of data collection. Therefore, we did not present the range of ANC contacts in this manuscript.

4. In addition, this study was conducted prior to the updated WHO recommendations of 8 ANC visits. However, it would be interesting to see the influence of this on the studies results or read the authors reflections on 8 visits at ANC as opposed to at least 4. The Ghana National Safe Motherhood Service protocol is from 2008 so naturally does include such updates.

- **Response to reviewer**

It would be interesting to see the feasibility of 8 ANC contacts in resource limited setting like rural area of Ghana. In our study sample, 19.8% had 8 ANC contacts. We assumed that these women did not necessarily have good care seeking behavior, but might have been in need of 8 contacts, such as to treat complications. In addition, 8 ANC contacts was out of scope in our intervention. Thus, we did not present the percentage of 8 ANC contacts in this paper.

Reviewer: 3

Reviewer Name: Khitam Muhsen

Institution and Country: Tel Aviv University, Israel

1. This is an interesting and important study. Please add to the abstract some details on the main statistical analysis. Please provide point estimates of the study outcomes at baseline and post-intervention in the abstract.
➢ **Response to reviewer** We added information on statistical analysis and presented point estimates of the study outcomes at the post-intervention. However, we could not present point estimates of the study outcomes at the baseline due to limited word count (L50-56)

2. Please rephrase the study aim according to PICO (population, intervention, comparison and outcome).
➢ **Response to reviewer** We revised the study aim in the abstract and the introduction. (L34-36, L116-121)

3. Please add more details on the control group. Did the control group receive the standard of care (what did it include?) or other intervention?
➢ **Response to reviewer** The control group received the conventional care recommended by the Ghana National Safe Motherhood Service Protocol. (L169-172)

4. It is not clear if the assumptions of the difference-in difference (DiD) method were met.
➢ **Response to reviewer** The DiD method is mainly required to meet two assumptions, according to “Impact evaluation in practice” authored by Gertler PJ, et. al. First, no time-varying difference exist between the intervention and the control groups. In this study, we did not observe any specific changes that might have affected the study outcome in both groups during the trial period. Second, the outcome trend should be equal in the intervention and the control groups in the absence of the trial. In this study, however, it was not feasible to measure the change that could have occurred in the intervention group in the absence of intervention, because we did not conduct any surveys before the baseline survey. Based on these considerations, we used the DiD method as the available method for our study. (L232-239)

5. Some characteristics of the intervention and control groups differed between the baseline and follow-up assessments, including parity, wealth quintiles and health insurance. It is worth assessing the impact of these factors, especially having health insurance on the study outcomes.
➢ **Response to reviewer** We agreed with the reviewer's comment. We adjusted basic characteristics of women in the difference-in-difference analysis (Tables 3 and 4) and the logistic regression analysis (Table 5) as follows: study site, living in a sub-district with a district hospital, age, education level, marital status, parity, religion, wealth quintiles, and the status of national health insurance membership. (L213-219).

6. Page 18, table 5: why this analysis was limited to the intervention group? Are the determinants of adequate contacts with high quality care different in the control group? An important message herein is reducing social disparities and increasing health insurance coverage to reduce gaps in adequate contacts with high quality care in the treatment of pregnant women and their babies.
➢ **Response to reviewer** We understand the reviewer's suggestion. However, we targeted the intervention group in the follow-up survey to identify the factors associated with the study outcome, because we wanted to identify barriers to having adequate contacts with high-quality care in the intervention group, and further improve implementation of the intervention package. (L254-257)

Reviewer: 4

Reviewer Name: Lifeng Lin

Institution and Country: Florida State University, USA

1. First, on page 7, the authors created 16 pairs from the 32 clusters, accounting for the population, the volume of delivery, and the number of midwife. I'm wondering if the pairing was within each of the 3 sites (Navrongo, Dodowa, and Kintampo). Also, were the clusters randomly paired? If yes, how was the randomization performed while the population, the volume of delivery, and the number of midwife needed to be similar for each pair of clusters? The authors may present a little more details about the pairing, e.g., providing cluster-level characteristics.
➢ **Response to reviewer** We made pairs of cluster within each of the 3 sites (6 pairs from Navrongo, 6 pairs from Kintampo, and 4 pairs from Dodowa). We paired the clusters with similar characteristics that could affect implementation and impact of the intervention, such as the population, the

volume of delivery, and the number of midwife. Then, we randomly assigned the two clusters within a pair to the intervention or the control groups. (L138-140)

2. Second, on page 12, the authors performed the logistic regression with several independent variables, including study site, parity etc. Were all independent variables treated as categorical variables? Was there any continuous variable? If all independent variables were categorical, the authors may consider presenting the detailed categories classified for each variable.

- **Response to reviewer**

  In this revision, the logistic regression model included the variables of study site, living in a sub-district with a district hospital, age, education, marital status, parity, religion, wealth quintile index, and the status of national health insurance membership as potential confounders. We treated all of these variables as categorical variables. We mentioned about this in the section of confounders. (L213-219)

3. Third, Table 2 on page 13 shows that religion and wealth quintiles were significantly different in the control and intervention groups, which may confound the final results. Although the authors mentioned that the intervention group had more Muslim and wealthy women than the control group, I think they may discuss more on these significant differences. What were the potential impact of these two variables on the effect of the CoC intervention? As wealthier women tended to receive better healthcare services, was wealth quintile possible a confounder? Also, was religion spatially related to the clusters' locations? These may be discussed as limitations in this last section.

- **Response to reviewer**

  We agree with the reviewer's point. Distributions of religion and socio-economic status were imbalanced between the intervention and the control groups. This could be because the distributions of these factors might be geographically related to clusters' locations as the reviewer mentioned. Therefore, we adjusted these individual characteristics in the revised analysis. (L213-219) In addition, we mentioned that "the clusters in the study were not homogeneous and cluster allocation was uneven. This might have impacted the effects of the intervention." in the limitation section. (L413-414)

4. On page 11, the second line from the bottom. Should "efficacious" be "efficacy"?

- **Response to reviewer**

  Yes, it should be "efficacy". We corrected. (L250)

5. On page 12, I think it is better to put the sentences at the top (for the per-protocol analysis, … excluded 238 women … and 134 women …) in the results section.

- **Response to reviewer**

We shifted the sentences to the result section (L307-308).

6. On page 12, the authors mentioned that they used “robust variance estimate” to control for potential correlations within clusters. I am not very clear what was the exact robust estimate used for variance and how it was related to the potential correlations. The authors may provide some citations on this or give a bit more details.

- **Response to reviewer**

The regression model has an assumption that the study sample is independent with each other. However, we used two-stage random sampling in the survey. In addition, we used DiD estimates to evaluated the effect of the intervention on the study outcome. That is, our sampled observations were collected under repeated measurement. These features could cause intra-class correlations in our models, and lead to inaccurate estimates of variances (i.e., standard errors, and confidence interval). Therefore, we used robust estimators of variance in our models to mitigate such features. We explained it in the manuscript as follows: Therefore, we performed the DiD analysis with cluster robust estimators of variance, controlling for individual characteristics. Robust estimators of variance is a technique used to estimate cluster robust standard errors and adjust the confidence intervals of the DiD estimators when the regression model is potentially affected by cluster correlations. (L238-242).

VERSION 2 – REVIEW

REVIEWER	Lifeng Lin Florida State University, USA
REVIEW RETURNED	20-Apr-2019

GENERAL COMMENTS	The authors have addressed my previous comments, and I do not have further suggestions from the statistical perspective.