PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The effects of a household conditional cash transfer programme on coverage and quality of antenatal care: a secondary analysis of Indonesia's pilot programme
AUTHORS	Triyana, Margaret; Shankar, Anuraj

VERSION 1 – REVIEW

REVIEWER	David ZOMBRE University of Montreal Public Health Research Institute, School of Public Health, Montreal, Québec, Canada Pavillon 7101 avenue du Parc C.P 6128 Succursale C, local 3224 Montréal, Québec, Canada H3C 3J7
REVIEW RETURNED	02-Nov-2016

GENERAL COMMENTS	This manuscript describes the effectiveness of household conditional cash transfer programmes (CCT) to increase the coverage of prenatal care and improve quality of prenatal care. The research question and study objective are clearly defined, but there are several major issues with the manuscript that need to be addressed. Major issues: *The outcomes are clearly defined as 1) coverage of prenatal care and 2) quality of prenatal care. 1. We have no details on how the items of each outcome were measured. 2. The authors said they used PCA but do not give enough information for us to appreciate the validity of their analyzes. (Bartlett's sphericity test and the KMO index.). 3. I have some doubts about the way THE outcomes were measured, since the authors have not told us if they use PCA with an already validated tool in the literature. How many factors have the authors found in their analysis? What components of coverage or quality of prenatal car could these factors be related? 4. The authors used simple PCA with dichotomous variables while polycohoricpca analysis is recommended when we have several dichotomous items. *The study design has not been well described. The authors have not clearly stated if they compare the two groups or if they compare the results before and after. *The authors must reconsider the results section because tables and results were not clearly presented.

	*The authors state that: “These results suggest that CCT program increasing was successful in the prenatal clinical services received by poor coverage households”, but we do not see in their manuscript how the authors compare the results of women by socio-economic status. In addition, the authors should not cite the results of other studies in theirs results section. *Finally, the result section has to be reconsidered to align the results with the objectives of the study.

REVIEWER	Tara Patricia Cookson Private Sector; Canada
REVIEW RETURNED	24-Jan-2017

GENERAL COMMENTS	The objective of this paper - to assess the effects of CCTs on quality of health care received by CCT recipients, indeed a relatively under-researched question, is very important. The article is accessible and very well written. On page 3, para 3 / page 12, para 2, you rightly identify a gap between utilization of health services and health outcomes. I would recommend a piece by Lena Lavinas [(2013) 21st Century Welfare, New Left Review (84)] which shows that while many states have implemented CCT programs and thus increased demand for services, the increase in demand has not been met with a proportionate investment in the supply side (ie investment in health and education services). Lavinas’ piece also speaks to the observation you make on page 13, line 45 regarding demand for services. The relationship between demand and supply is crucial context for your findings, which suggest that getting poor people to use services does not automatically equate with a reduction in poverty. Qualitative research has also underscored the gap between high rates of compliance with conditions and low quality of services, and relatedly, the physical cost and time burden borne by women travelling to and from under-resourced clinics (that suffer frequent clinic closures, understaffing, and stockouts), particularly in rural areas. My article on the Peruvian CCT is an example: Cookson, T.P. Working for Inclusion? Conditional Cash Transfers, Rural Women and the Reproduction of Inequality. Antipode, 38(5). Quantitative and qualitative research read together makes a compelling case for improving service quality. I am curious about the distinction between ‘coverage’ (which I understand from your article to be services provided), reported on by expectant women, and ‘quality’, reported on by midwives. Why ask midwives to report on quality? Was there an opportunity to gather or analyze data on expectant women’s perception of service quality (although I understand that in using either public or private services, this would require a distinct sort of survey)? Certainly in Peru, there is a difference in what rural, indigenous women and urban, mestizo health professionals understand to be quality care (in this case, relating largely to discriminatory attitudes towards poor women and preferences relating to style of birth). Perhaps on page 5 para 2, there could be an additional brief definition of ‘quality’, noting that this excludes social aspects.

	The difference in coverage between public and private practice is intriguing (page 12 para 1). Why is this? Why do some women access private clinics and others public? It might be worth highlighting this as an area for future research. I hope that these comments – many of which come from a qualitative orientation – are useful to helping your findings reach the broadest of audiences possible. Your research question is of great interest to qualitative researchers of CCTs, and I do hope that your findings get brought into qualitative discussions!

REVIEWER	Dr. Avishek Hazra Population Council, India
REVIEW RETURNED	30-Jan-2017

GENERAL COMMENTS	Broad comments: 1. The article talks about an important aspect of quality of care during prenatal period and its association with CCT. However, the manuscript may be strengthened in certain areas. For example, the RCT design and the CCT program needs more elaboration for an external reader. It fails to provide how the randomization was done, was there any specific age specification of children for their mothers to be eligible for the CCT, how the amount was dependent on 'household composition' and whether beneficiaries were eligible to receive the amount upon meeting all the requirements/conditions or at least one condition. 2. The study design and methods section is weak and demands for improvement. The paper does not mention about the eligibility criteria of sample to be included in the study, how the 2,723 villages were selected for the survey, whether the selected villages were from all the 329 treatment and 259 control sub-districts or from a sample of these sub-districts. Although the paper gave a reference of impact evaluation (14), this paper must be independent in its own and provide reader at least the basic information. The year of evaluation waves should be specified in the data source section, not in the population section. It says baseline and follow-up surveys included 'pregnancy history in the 24 months prior to the survey'. It is not at all clear if the respondents included currently pregnant as well as women delivered in last 24 months; if the later one, whether the index pregnancy was the latest one or the previous one in case of more than one pregnancy in last 2 years. 3. The analysis presented in Table 2 seems problematic. What are the reference categories? Are the changes over time or in treatment group over control group or net changes? Did the authors take time (baseline/follow-up) and type (treatment/control), or the interaction term as independent variables? Else, how does the authors prove that the increased odds are due to program impact? 4. Analysis of Table 3 is also not convincing. In page 5 of the paper it is mentioned that "The quality of prenatal care provided was only asked in the follow-up survey, so the analysis was based on cross-sectional data." Then what is meant by this sentence - "Table 3 presents changes in prenatal quality among dual practice midwives" as written in page 13, line 27?

	Specific comments: 1. Abstract: a. It is not clear how many years after the CCT intervention the evaluation was conducted? The authors may like to add the duration of CCT and year of conducting the impact evaluation survey-baseline and follow-up. b. In the results section, the reference category is not mentioned and puts the reader in confusion. For example, "Women were more likely to receive the following services…", but unless the reference category is mentioned, it is unclear whether more likely than control group or more likely in time-2 than time-1 (assuming it's a pre-post experimental-control design, which should be clearly mentioned in the Setting section) c. In the conclusion section it is written "…but this change was not associated with improved pregnancy outcomes partly because midwives did not improved PNC quality". Suggest authors to add a line about this in the result section. 2. Main text a. Page 3, Line 39: It says "In spite of this puzzle…". It is not very clear about what is the puzzle the authors referring to. The statement needs explanation. b. Page 3, Line 58: "There were 329 sub-districts randomised into treatment and 259 sub-districts randomised into the control group". Readers may be curious to know how the randomization was done, whether pairs of sub-districts were created, if yes, based on what criteria etc. c. Page 4, line 10: it says cash transfers to expectant women and mothers. It needs to be specified if authors refer to mothers of expectant women or mothers with children of specific age group? d. Page 4, line 12-13: "The transfer amounted to 15 to 20% of estimated total consumption of poor households." It is unclear whether total monthly consumption or quarterly consumption or yearly consumption? To whom the cash transfers were made – to head of the households or to the expectant women and mothers with children of specific age? e. Page 4, under data source, please specify who the respondents were or what the eligibility criteria was. f. Page 4, under variables section it says "The prenatal care items included the following dichotomous variables: measurements of women's weight, height, blood pressure, fundal height, and fetal heartbeat. In addition, women should also receive a blood test (for syphilis and HIV), external and internal pelvic examinations, 90 iron pills, two tetanus toxoid vaccinations, information on signs of pregnancy complications, and what to do if there were signs of pregnancy complications. Please indicate if indicators mentioned after 'in addition' (such as blood test, pelvic exam etc.) were also included in the principal component analysis. g. Table 1: The values corresponding to per capita household expenditure needs to be re-checked as mean and SD in treatment and control are identical, and also the adjusted difference values looks absurd. h. Table 2: The interpretation needs to clearer. As pointed out earlier in 1b, whenever authors talk about 'more likely or less likely' they should specify the reference category. For example, "Expectant women living in treated communities were more likely than xxxxxxx to receive the following services during pregnancy…" Also, in the table it is mentioned 'wave 3'. Please clarify what it is, as not mentioned anywhere in the article.

	i. Table 3: Every time, please specify the reference categories while interpreting odds ratios. Finally, I would recommend the authors to look at some of the papers that describes analysis and results of logistic regressions, and revise this paper accordingly. Further, just by analysing one-point data (follow-up in this case) of health care providers, concluding that program had no significant effect on quality of care may not be true always.

REVIEWER	Sarika Chaturvedi Savitribai Phule Pune University, India I have no competing interests.
REVIEW RETURNED	09-Feb-2017

GENERAL COMMENTS	This is a highly relevant work. The authors have used appropriate methods for the study. The manuscript needs improvement on the following lines: 1. The manuscript requires English language revision.Currently it does not stand for an international journal, especially the abstract. 2. Abstract- Outcome measures- prenatal services provided by midwives seems - and also reported by midwives, this could be confusing- pl clarify like midwives reports of prenatal services provided . 3.Abstract- Is important to include after how many years of CCT implementation does the evaluation refer to. 4.Introduction- Line 32- Nine is misspelt at 'nice'. 5. About the PKH- In the pilot attendance sheets are mentioned to be collected from schools, this is not clear. 6. Discussion- pg 14-Line 9-11: Authors refer to the results of Barber et al, however from the writing it appears they refer to their own results. The statement needs to be revised for clarity. 7. The comparison of care between public and private sectors is interesting. It would be important to discuss this finding in detail. The authors may add in the background info about access to private sector among the target population for PHK. 8. The results are from the 2009 survey, albeit old for a manuscript in 2017. However given the importance of the analysis done, the findings may be considered new; it would be important that the authors discuss the findings in the present context- what has changed in the Indonesian health system/PKH since 2009 and how they would interpret these findings in that light. 9. The authors point to issues with accountability, this discussion needs to be elaborated. 10. Also, the impact of CCTs on quality of services needs to be discussed in details with reference to literature from other countries too. This pattern seems to be global, the authors may discuss this trend. What are the underlying problems with the assumptions of CCTs and the related equipedness issues.

REVIEWER	Ngianga II Kandala University of Portsmouth, United Kingdom None
REVIEW RETURNED	27-Mar-2017

GENERAL COMMENTS	The paper is clear, reads well, appropriate statistical methods are used. The reviewer also provided a marked copy with additional comments. Please contact the publisher for full details.

REVIEWER	Edwin Amalraj Raja Medical Statistics Team, University of Aberdeen, Aberdeen, UK None
REVIEW RETURNED	12-Apr-2017

GENERAL COMMENTS	The manuscript on 'The effect of Household conditional cash transfer (CCT) on coverage and quality of prenatal care: Evidence from Indonesia" by Triyana et al reads well. The clinical importance of the manuscript will be evaluated by others in the specialty. However, I have a few points which require attention and recommend that these need to be taken into account for revision: 1. This study's objective is to find effect of CCT on prenatal coverage reported by eligible women and on prenatal provider quality reported by midwives. It is not very clear whether the individual prenatal care items are outcomes of the study or only the PNC index derived from the principal component analysis using the prenatal care items. It is also the same for PNC quality items and also to PNC quality index. 2. It is mentioned that PNC coverage or PNC provider care items are binary. There is a duplicate of the same information in 'variable & covariate' and in 'Statistical analysis'. The author should mention very clearly whether the items have clinically meaningful cut off and if so, what value of the item above or below is considered to be reference/control category. Or whether receiving (PNC self-reported by women) the items indicates just the measurement of the item. i.e., the response was yes for that item. Was there 'No'? Can you consider missing as 'No'? 3. The statement 'we used least square regression….' is very vogue. The author should mention what is the purpose of using least square regression, dependent and independent variables (and also for logistic regression). 4. The total number of women in treatment and control group does not add upto total pregnancies reported in figure 1. 5. One p-value can be presented for each variable (table 1) to indicate statistical significance of the association. There seems to be more than 2 p-values for some variables which were having more than two categories. 6. The table for comparison of baseline characteristics between treatment and control may be presented with number & percent rather than mean & SD. The author would have used 'chi-square test / t-test for clustered

	data' to adjust for sub-district level clustering of study subjects or use the p-value from the standard univariate model (including sub-district as indicator variable). 7. Kindly check numbers related to 'Per capita household expenditure' and 'male child' in the table 1. 8. The statement 'the majority of women in the sample were under 30 in 2007' does not seem to be correct and not verifiable from the table. 9. It is mentioned in the methods 'All reported pregnancies in both waves of the survey were included in the analysed sample' and also in the 1st paragraph of study population. But table 2 provides results of wave 3. Please clarify. 10. Using pca to derive an index based on many correlated factors/items for each individual is a novel idea. It seems the authors have derived index/score thro' one component. On what basis did you decide? Are there more than one component in the pca? There was no mention about what is the variance explained by this component or to the rest of the other components. This will relate to the reliability of the index/score. 11. Missing- indicator method is used to do complete case analysis but has been criticized for introducing bias. Is there any alternatives? (Greenland S et al. A critical look at methods for handling missing covariates in epidemiologic regression analyses. AJE 1995; 142(12): 1255-1264) 12. Please present OR and its 95% CI rounding to 2 decimal precision. Further when the 95% CI does not include 1 the association is statistically significant. There may be a statement relating CI and p-value in methods and avoid presenting p-value along with CIs (in the text). 13. In table 2, what is the summary measure of 'tetanus vaccinations'? The summary measure indicates negative but its lower CI does not. 14. The software has limitation in showing actual p-value in the output. The p-value of 0.000 means p-value < 0.001. So please don't present 0.000 as p-value. 15. The column headings may have 'OR' and '95% CI' for the respective columns in table 2 & 3. If the statistical measure is different from OR, please indicate that measure in footnote using a special mark. 16. The table 2 presents the effect of CCT on PNC items and index for pooled cases (baseline + follow up) and separately for private and public from follow up survey. Since the results of pooled data analysis is different from practice type it is recommended to present ' pooled results of the follow up survey' and to see whether the practice type is an effect modifier. You may also include pooled results (baseline + follow up) and explain the differences, if any, between pooled & follow up survey. Has any attempt made to find whether the coverage or quality changed from baseline to follow up? You may use either 'fixed effect' modelling under multilevel frame work or use time * CCT interaction in the longitudinal analysis.

	17. Table 3 for the results of PNC quality may be revised as suggested for table 2 for the results of PNC coverage. 18. The statement 'while a longitudinal … results' in the study population and ' the lack of …..over time' in discussion did not clearly specify to the readers what was available and what was not in order to conduct longitudinal data analysis. 19. It is not practice to interpret the results and put references in the result section. 20. The paragraph under PNC quality in the results section did not contain any results. The author should move the paragraph from results section & put in an appropriate place. 21. Use the word either 'women' or 'patient' to indicate study subject in the manuscript The manuscript requires a major revision with re-analysis of data

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: David ZOMBRE
Institution and Country: University of Montreal Public Health Research Institute, School of Public Health, Montreal, Québec, Canada, Pavillon 7101 avenue du Parc, C.P 6128 Succursale C, local 3224, Montréal, Québec, Canada, H3C 3J7
Please state any competing interests: None declared

Please leave your comments for the authors below

This manuscript describes the effectiveness of household conditional cash transfer programmes (CCT) to increase the coverage of prenatal care and improve quality of prenatal care.
The research question and study objective are clearly defined, but there are several major issues with the manuscript that need to be addressed.

Major issues:
*The outcomes are clearly defined as 1) coverage of prenatal care and 2) quality of prenatal care.
1. We have no details on how the items of each outcome were measured.
Response: We have clarified the variables included in the PCA in the 'Variables and covariates' section.

2. The authors said they used PCA but do not give enough information for us to appreciate the validity of their analyzes. (Bartlett's sphericity test and the KMO index.).

Response: Thank you for the suggestion. We included the Bartlett's sphericity test and the KMO index in the 'Methods' section. The results for the household survey are: Bartlett's sphericity test p-value < 0.001 and KMO index 0.736. The results for the midwives are: Bartlett's sphericity test p-value < 0.001 and KMO index 0.796. These suggest that PCA is appropriate in this case.
The items included in the PCA are based on the guidelines for required care provided by the Indonesian Ministry of Health. We have added this reference in the 'Variables and covariates' section.
The PCA for women resulted in 3 components with eigenvalues greater than 1.

We selected the primary component which accounted for 61% of the variance. While the index using PCA was a primary outcome, we also analyzed individual items. Taken together, these results provide evidence for the impact of the programme on overall PNC and specific components.

3. I have some doubts about the way THE outcomes were measured, since the authors have not told us if they use PCA with an already validated tool in the literature. How many factors have the authors found in their analysis? What components of coverage or quality of prenatal car could these factors be related?

Response: The items included in the PCA are based on the guidelines for required care provided by the Indonesian Ministry of Health. We have added this reference in the 'Variables and covariates' section. We selected the primary component for the women and on the supply side, for public and private practice.

4. The authors used simple PCA with dichotomous variables while polycohoricpca analysis is recommended when we have several dichotomous items.
Response: We have included the additional analysis and discuss the results. They are qualitatively similar to the results using the 'pca' command.

*The study design has not been well described. The authors have not clearly stated if they compare the two groups or if they compare the results before and after.

Response: We clarified the analyses in Tables 2 and 3 by including additional notes. Table 2 uses the pooled sample (baseline and follow-up) and the cross-sectional analysis (public vs. private). Table 3 uses cross-sectional analysis.
*The authors must reconsider the results section because tables and results were not clearly presented.

Response: We have added notes to Tables 2 and 3.
*The authors state that: "These results suggest that CCT program increasing was successful in the prenatal clinical services received by poor coverage households", but we do not see in their manuscript how the authors compare the results of women by socio-economic status. In addition, the authors should not cite the results of other studies in theirs results section.

Response: We included this in the 'Study design and data source' section. The sample includes near poor and poor households, so the results apply to the population that is targeted by the program.

*Finally, the result section has to be reconsidered to align the results with the objectives of the study.
Response: We have linked the results from the household survey and the provider survey to highlight that the improvements reported by households are driven by utilization.

Reviewer: 2
Reviewer Name: Tara Patricia Cookson
Institution and Country: Private Sector; Canada
Please state any competing interests: None declared

Please leave your comments for the authors below

The objective of this paper - to assess the effects of CCTs on quality of health care received by CCT recipients, indeed a relatively under-researched question, is very important.

The article is accessible and very well written.

Comment: On page 3, para 3 / page 12, para 2, you rightly identify a gap between utilization of health services and health outcomes. I would recommend a piece by Lena Lavinas [(2013) 21st Century Welfare, New Left Review (84)] which shows that while many states have implemented CCT programs and thus increased demand for services, the increase in demand has not been met with a proportionate investment in the supply side (ie investment in health and education services). Lavinas' piece also speaks to the observation you make on page 13, line 45 regarding demand for services. The relationship between demand and supply is crucial context for your findings, which suggest that getting poor people to use services does not automatically equate with a reduction in poverty.

Response: We have included the reference, thank you for pointing this out.

Comment: Qualitative research has also underscored the gap between high rates of compliance with conditions and low quality of services, and relatedly, the physical cost and time burden borne by women travelling to and from under-resourced clinics (that suffer frequent clinic closures, understaffing, and stockouts), particularly in rural areas. My article on the Peruvian CCT is an example: Cookson, T.P. Working for Inclusion? Conditional Cash Transfers, Rural Women and the Reproduction of Inequality. Antipode, 38(5). Quantitative and qualitative research read together makes a compelling case for improving service quality.

Response: We have included the reference, thank you for pointing this out.

Comment: I am curious about the distinction between 'coverage' (which I understand from your article to be services provided), reported on by expectant women, and 'quality', reported on by midwives. Why ask midwives to report on quality? Was there an opportunity to gather or analyze data on expectant women's perception of service quality (although I understand that in using either public or private services, this would require a distinct sort of survey)? Certainly in Peru, there is a difference in what rural, indigenous women and urban, mestizo health professionals understand to be quality care (in this case, relating largely to discriminatory attitudes towards poor women and preferences relating to style of birth). Perhaps on page 5 para 2, there could be an additional brief definition of 'quality', noting that this excludes social aspects.

Response: We unfortunately do not have additional data to answer these interesting questions. We have included this in the 'Variables and covariates' section.

Comment: The difference in coverage between public and private practice is intriguing (page 12 para 1). Why is this? Why do some women access private clinics and others public? It might be worth highlighting this as an area for future research.

Response: We are unfortunately unable to model why women selected private or public practice. We have included this limitation in our discussion of the results.

Comment: I hope that these comments – many of which come from a qualitative orientation – are useful to helping your findings reach the broadest of audiences possible. Your research question is of great interest to qualitative researchers of CCTs, and I do hope that your findings get brought into qualitative discussions!

Reviewer: 3
Reviewer Name: Dr. Avishek Hazra
Institution and Country: Population Council, India
Please state any competing interests: 'None declared'

Please leave your comments for the authors below

Please refer to my comments, attached

Broad comments:
1.The article talks about an important aspect of quality of care during prenatal period and its association with CCT. However, the manuscript may be strengthened in certain areas. For example, the RCT design and the CCT program needs more elaboration for an external reader.

Response: We have clarified the design in the 'Study design and data source' section.

Comment: It fails to provide how the randomization was done, was there any specific age specification of
children for their mothers to be eligible for the CCT, how the amount was dependent on 'household composition' and whether beneficiaries were eligible to receive the amount upon meeting all the requirements/conditions or at least one condition.

Response: We unfortunately do not have data on compliance to each requirement. However, we include in the manuscript that anecdotally, households continue to receive the transfers so long as they meet at least 1 program requirement. We have clarified the eligibility criteria and transfer amounts for pregnant women in the 'Study design and data source' section.

2. The study design and methods section is weak and demands for improvement. The paper does not mention about the eligibility criteria of sample to be included in the study, how the 2,723 villages were selected for the survey, whether the selected villages were from all the 329 treatment and 259 control sub-districts or from a sample of these sub-districts. Although the paper gave a reference of impact evaluation (14), this paper must be independent in its own and provide reader at least the basic information. The year of evaluation waves should be specified in the data source section, not in the population section. It says baseline and follow-up surveys included 'pregnancy history in the 24 months prior to the survey'. It is not at all clear if the respondents included currently pregnant as well as women delivered in last 24 months; if the later one, whether the index pregnancy was the latest one or the previous one in case of more than one pregnancy in last 2 years.

Response: We have moved the description of the survey to the 'Study design and data source' section. We have included a description of the sample selection and clarified that the households are near poor and poor. We also include the pregnancies in the analyzed sample.

3. The analysis presented in Table 2 seems problematic. What are the reference categories? Are the changes over time or in treatment group over control group or net changes? Did the authors take time (baseline/follow-up) and type (treatment/control), or the interaction term as independent variables? Else, how does the authors prove that the increased odds are due to program impact?

Response: We have clarified the pooled vs cross-sectional analysis by including notes to table 2. We compared outcomes for women residing in control and treated areas, with district fixed effects to take into account non time-varying district characteristics. The coefficient would then be the average difference in outcomes for women residing in treated areas compared to the control. We have included this in the 'Statistical analysis' section.

4. Analysis of Table 3 is also not convincing. In page 5 of the paper it is mentioned that "The quality of prenatal care provided was only asked in the follow-up survey, so the analysis was based on cross-sectional data." Then what is meant by this sentence - "Table 3 presents changes in prenatal quality among dual practice midwives" as written in page 13, line 27?

Response: Thank you for pointing that out. We meant the 'difference' since we could only do cross sectional analysis for midwife quality. However, we compared the baseline characteristics of midwives and they are similar, so the coefficient can be interpreted as program impact. We believe that the cross sectional analysis remains relevant in spite of the data limitation.

Specific comments:
1. Abstract:
a. It is not clear how many years after the CCT intervention the evaluation was conducted? The authors may like to add the duration of CCT and year of conducting the impact evaluation survey-baseline and follow-up.

Response: We have included the timing and survey years in the abstract.

b. In the results section, the reference category is not mentioned and puts the reader in confusion. For example, "Women were more likely to receive the following services…", but unless the reference category is mentioned, it is unclear whether more likely than control group or more likely in time-2 than time-1 (assuming it's a pre-post experimental-control design, which should be clearly mentioned in the Setting section)

Response: We have clarified this in the results section.

c. In the conclusion section it is written "…but this change was not associated with improved pregnancy outcomes partly because midwives did not improved PNC quality". Suggest authors to add a line about this in the result section.

Response: We have changed this, thank you.

2. Main text
a. Page 3, Line 39: It says "In spite of this puzzle…". It is not very clear about what is the puzzle the authors referring to. The statement needs explanation.

Response: We have rewritten the introduction.

b. Page 3, Line 58: "There were 329 sub-districts randomised into treatment and 259 subdistricts randomised into the control group". Readers may be curious to know how the randomization was done, whether pairs of sub-districts were created, if yes, based on what criteria etc.

Response: We included a description of the sample selection in the 'Study design and data source' section.

c. Page 4, line 10: it says cash transfers to expectant women and mothers. It needs to be specified if authors refer to mothers of expectant women or mothers with children of specific age group?

Response: We have clarified this in the 'Study design and data source' section.

d. Page 4, line 12-13: "The transfer amounted to 15 to 20% of estimated total consumption of poor households." It is unclear whether total monthly consumption or quarterly consumption or yearly consumption? To whom the cash transfers were made – to head of the households or to the expectant women and mothers with children of specific age?

Response: We have clarified this in the 'Study design and data source' section.

e. Page 4, under data source, please specify who the respondents were or what the eligibility criteria was.

Response: We have clarified this in the 'Study design and data source' section.

f. Page 4, under variables section it says "The prenatal care items included the following dichotomous variables: measurements of women's weight, height, blood pressure, fundal height, and fetal heartbeat. In addition, women should also receive a blood test (for syphilis and HIV), external and internal pelvic examinations, 90 iron pills, two tetanus toxoid vaccinations, information on signs of pregnancy complications, and what to do if there were signs of pregnancy complications. Please indicate if indicators mentioned after 'in addition' (such as blood test, pelvic exam etc.) were also included in the principal component analysis.

Response: We have clarified this in the 'Variables and covariates' section.

g. Table 1: The values corresponding to per capita household expenditure needs to be rechecked as mean and SD in treatment and control are identical, and also the adjusted difference values looks absurd.

Response: Thank you for catching this error. We have corrected the 95% CI in Table 1. The adjusted difference estimates are consistent with the initial impact evaluation.

h. Table 2: The interpretation needs to clearer. As pointed out earlier in 1b, whenever authors talk about 'more likely or less likely' they should specify the reference category. For example, "Expectant women living in treated communities were more likely than xxxxxxx to receive the following services during pregnancy…" Also, in the table it is mentioned 'wave 3'. Please clarify what it is, as not mentioned anywhere in the article.

Response: We have included this in the results section.

Comment: Table 3: Every time, please specify the reference categories while interpreting odds ratios.

Response: We have included this.

Comment: Finally, I would recommend the authors to look at some of the papers that describes analysis and results of logistic regressions, and revise this paper accordingly. Further, just by analysing one-point data (follow-up in this case) of health care providers, concluding that program had no significant effect on
quality of care may not be true always.

Response: We include the data limitation in our discussion.

Reviewer: 4
Reviewer Name: Sarika Chaturvedi
Institution and Country: Savitribai Phule Pune University, India
Please state any competing interests: I have no competing interests.

Please leave your comments for the authors below

This is a highly relevant work. The authors have used appropriate methods for the study. The manuscript needs improvement on the following lines:

1. The manuscript requires English language revision.Currently it does not stand for an international journal, especially the abstract.

Response: We have edited the manuscript.

2. Abstract- Outcome measures- prenatal services provided by midwives seems - and also reported by midwives, this could be confusing- pl clarify like midwives reports of prenatal services provided .

Response: We have clarified the self-reported measures.

3.Abstract- Is important to include after how many years of CCT implementation does the evaluation refer to.

Response: We have included this in the ‘Methods’ section.

4.Introduction- Line 32- Nine is misspelt at 'nice'.

Response: We have corrected this, thank you for pointing it out.

5. About the PKH- In the pilot attendance sheets are mentioned to be collected from schools, this is not clear.

Response: We have removed the education requirements from manuscript to avoid this confusion.

6. Discussion- pg 14-Line 9-11: Authors refer to the results of Barber et al, however from the writing it appears they refer to their own results. The statement needs to be revised for clarity.

Response: Our results are in line with earlier results, we have rewritten this.

7. The comparison of care between public and private sectors is interesting. It would be important to discuss this finding in detail. The authors may add in the background info about access to private sector among the target population for PHK.

Response: We unfortunately do not have sufficient information on how/why women would choose private over public practice, so we have added this as another limitation to our study.

8. The results are from the 2009 survey, albeit old for a manuscript in 2017. However given the importance of the analysis done, the findings may be considered new; it would be important that the authors discuss the findings in the present context- what has changed in the Indonesian health system/PKH since 2009 and how they would interpret these findings in that light.

Response: We have included the expansion of PKH and the introduction of universal coverage in the results section to highlight the policy relevance of our study.

9. The authors point to issues with accountability, this discussion needs to be elaborated.

Response: We could only limit our point to accountability in terms of the provision of specific PNC items, we have added this in the discussion.

10. Also, the impact of CCTs on quality of services needs to be discussed in details with reference to literature from other countries too. This pattern seems to be global, the authors may discuss this trend. What are the underlying problems with the assumptions of CCTs and the related equipedness issues.

Response: We have added this in the discussion section.

Reviewer: 5
Reviewer Name: Ngianga II Kandala
Institution and Country: University of Portsmouth, United Kingdom
Please state any competing interests: None

Please leave your comments for the authors below

Good job!

The paper is clear, reads well, appropriate statistical methods are used.
However, minor comments as in the attached should be addressed.

Comment: Page 5: This should not serve as evidence of causal links

Response: We have added the limitation of the cross-sectional analysis.

Comment: Page 7: The hierarchal nature of the data is obvious, patient, midwife and sub-district levels. I wonder why were Random effect models not used?

This would have enable you to quantify not only random effects(differences between sub-district) due to sub-district level , but possibly differences between midwife (midfiwe randon effects).

Response: While the random effects model would allow for midwife-random effects, we are concerned that there may be omitted variables that may be correlated with the explanatory variables. In this instance, fixed effects model would provide more conservative estimates of the effects.

Reviewer: 6

Reviewer Name: Edwin Amalraj Raja
Institution and Country: Medical Statistics Team, University of Aberdeen, Aberdeen, UK
Please state any competing interests: None

Please leave your comments for the authors below

The manuscript on 'The effect of Household conditional cash transfer (CCT) on coverage and quality of prenatal care: Evidence from Indonesia" by Triyana et al reads well. The clinical importance of the manuscript will be evaluated by others in the specialty.

However, I have a few points which require attention and recommend that these need to be taken into account for revision:

1. This study's objective is to find effect of CCT on prenatal coverage reported by eligible women and on prenatal provider quality reported by midwives. It is not very clear whether the individual prenatal care items are outcomes of the study or only the PNC index derived from the principal component analysis using the prenatal care items. It is also the same for PNC quality items and also to PNC quality index.

Response: We include the quality index and individual items to provide additional evidence of the impact of the CCT.

2. It is mentioned that PNC coverage or PNC provider care items are binary. There is a duplicate of the same information in 'variable & covariate' and in 'Statistical analysis'. The author should mention very clearly whether the items have clinically meaningful cut off and if so, what value of the item above or below is considered to be reference/control category. Or whether receiving (PNC self-reported by women) the items indicates just the measurement of the item. i.e., the response was yes for that item. Was there 'No'? Can you consider missing as 'No'?

Response: We use binary outcomes to construct the index. The missing responses were excluded to avoid imputation.

3. The statement 'we used least square regression….' is very vogue. The author should mention what is the purpose of using least square regression, dependent and independent variables (and also for logistic regression).

Response: We include the outcomes in the 'Statistical analysis' section.

4. The total number of women in treatment and control group does not add upto total pregnancies reported in figure 1.

Response: The number of observations in table 1 corresponds to pregnancies at baseline. The number of observations in the pooled analysis in table 2 corresponds to the total number of pregnancies in figure 1. The number of pregnancies in the cross sectional analysis in table 2 can only be classified into public or private when this variable is non-missing.

5. One p-value can be presented for each variable (table 1) to indicate statistical significance of the association. There seems to be more than 2 p-values for some variables which were having more than two categories.

Response: We have corrected table 1, thank you for pointing this out.

6. The table for comparison of baseline characteristics between treatment and control may be presented with number & percent rather than mean & SD. The author would have used 'chi-square test / t-test for clustered data' to adjust for sub-district level clustering of study subjects or use the p-value from the standard univariate model (including sub-district as indicator variable).

Response: Thank you, we've changed table 1 to percentages. We used a regression to adjust for sub-district clustering, and reported the 95% CI of the coefficients in table 1.

7. Kindly check numbers related to 'Per capita household expenditure' and 'male child' in the table 1.

Response: We have corrected the entry, thank you for pointing this out.

8. The statement 'the majority of women in the sample were under 30 in 2007' does not seem to be correct and not verifiable from the table.

Response: While we do not include the average age in the table, more than 50% of the sample at baseline were in the under 25 and 26-30 categories.

9. It is mentioned in the methods 'All reported pregnancies in both waves of the survey were included in the analysed sample' and also in the 1st paragraph of study population. But table 2 provides results of wave 3. Please clarify.

Response: This is indeed confusing. We used the pooled analysis for the average change, but only used the follow-up survey to analyse private versus public practice. We have included an explanation in the notes to table 2 and in the results section.

10. Using pca to derive an index based on many correlated factors/items for each individual is a novel idea. It seems the authors have derived index/score thro' one component. On what basis did you decide? Are there more than one component in the pca? There was no mention about what is the variance explained by this component or to the rest of the other components. This will relate to the reliability of the index/score.

Response: We have added robustness to the use of pca to address your concern and also included the Bartlett's sphericity test and KMO index to test the appropriateness of pca in the sample. We have also added information on the number of components generated with eigenvalues greater than 1, and the variance explained by the primary component.

11. Missing- indicator method is used to do complete case analysis but has been criticized for introducing bias. Is there any alternatives? (Greenland S et al. A critical look at methods for handling missing covariates in epidemiologic regression analyses. AJE 1995; 142(12): 1255-1264)

Response: We decided against imputing missing indicators for this analysis. The number of cases with missing data were x% and analysis of distribution of covariates revealed little bias, suggesting data were missing at random. We therefore felt the methods used would still yield reliable results at the population level.

12. Please present OR and its 95% CI rounding to 2 decimal precision. Further when the 95% CI does not include 1 the association is statistically significant. There may be a statement relating CI and p-value in methods and avoid presenting p-value along with CIs (in the text).

Response: We followed another article in using 3 decimal points, but we can revise the tables to 2 decimal points. We have removed the p-values from the tables and only included p-values in the text.

13. In table 2, what is the summary measure of 'tetanus vaccinations'? The summary measure indicates negative but its lower CI does not.

Response:Thank you for pointing this out, we have corrected this.

14. The software has limitation in showing actual p-value in the output. The p-value of 0.000 means p-value < 0.001. So please don't present 0.000 as p-value.

Response: Thank you for pointing this out, we have corrected this.

15. The column headings may have 'OR' and '95% CI' for the respective columns in table 2 & 3. If the statistical measure is different from OR, please indicate that measure in footnote using a special mark.

Response: We have added this, thank you.

16. The table 2 presents the effect of CCT on PNC items and index for pooled cases (baseline + follow up) and separately for private and public from follow up survey. Since the results of pooled data analysis is different from practice type it is recommended to present ' pooled results of the follow up survey' and to see whether the practice type is an effect modifier. You may also include pooled results (baseline + follow up) and explain the differences, if any, between pooled & follow up survey. Has any attempt made to find whether the coverage or quality changed from baseline to follow up? You may use either 'fixed effect' modelling under multilevel frame work or use time * CCT interaction in the longitudinal analysis.

Response: We unfortunately do not have additional data to be able to conduct more detailed analysis to address this important question. Our cross-sectional analysis does not answer the question on how/why women whose private over public practice. We include this as a limitation to our study in the results section. The differences in quality as reported by midwives also do not present quality changes due to the cross-section analysis.

17. Table 3 for the results of PNC quality may be revised as suggested for table 2 for the results of PNC coverage.

Response: We have formatted table 3 similarly as you suggested.

18. The statement 'while a longitudinal … results' in the study population and ' the lack of …..over time' in discussion did not clearly specify to the readers what was available and what was not in order to conduct longitudinal data analysis.

Response: We have changed this to note that only a cross-section was available for some of the analyses.

19. It is not practice to interpret the results and put references in the result section.

Response: We have changed this.

20. The paragraph under PNC quality in the results section did not contain any results. The author should move the paragraph from results section & put in an appropriate place.

Response: We have moved this.

21. Use the word either ‘women’ or ‘patient’ to indicate study subject in the manuscript

Response: We have changed this to ‘women’ throughout, thank you for pointing this out.
The manuscript requires a major revision with re-analysis of data
We have included additional analysis

VERSION 2 – REVIEW

REVIEWER	David ZOMBRE University of Montreal School of Public Health, Canada
REVIEW RETURNED	17-Jun-2017

GENERAL COMMENTS	The authors made very useful and appropriate modifications to the previous version. I have no additional suggestions on the manuscript at this time.

REVIEWER	Tara Patricia Cookson University of British Columbia, Canada
REVIEW RETURNED	21-Jun-2017

GENERAL COMMENTS	The authors made significant improvements to the paper and I recommend it for publication contingent upon a few very minor changes, listed below. Pg 5 Line 15 grammatical error: "we includes" --> "we include" Pg 12 line 10 grammatical error: "This low proportion is may be due" --> "This low proportion may be due" Pg 12 lines 9 - 11: You suggest that only 20% of women had pelvic exams perhaps due to "the infrastructure of the healthcare facility and cultural norms". This seems insufficient, and while your study could not delve into the other material and social determinants of healthcare quality, it is important to at least consider and preferably point to research that considers the reasons why women receive inadequate sexual and reproductive healthcare. "Infrastructure" is too broad to pin women not receiving important pelvic exams. If you can, point to research that does explore this within the Philippine context. Pg 12 lines 16-20: You suggest that women not receiving the full schedule of iron pills is likely due to poor compliance with PNC visits and iron supplementation. How do you draw this correlation?

	How do you know it was not because of stock-outs? And how do you know that following this compliance with PNC visits became a requirement? Pg 15 lines 50-53: I am not sure what you mean in these two sentences regarding price dimension versus quality dimension. In the following sentence you (correctly) identify that when incentives are only provided to service users and not service providers, there may not be a motivation to improve service quality. But does this contradict what you say before? This paragraph needs some clarification work.

REVIEWER	Dr. Avishek Hazra Population Council, India
REVIEW RETURNED	18-Jun-2017

GENERAL COMMENTS	The article now reads much better than the previous version. However, below are few suggestions that the authors might consider for further improvement. 1. I have strong reservation in the statement that “the study is exploring the link between PNC component coverage for specific service items and PNC provider quality of midwives” (as mentioned in line 40-41 in page 3). The study has independently shown the effect of the program on PNC utilization and provider’s quality using two separate datasets. Nowhere in the link between the two has been established through data. Exploring the link is only possible if individual women data could be linked/merged/appended to corresponding provider data, and then assigning average score for each outcome to all individual women who are linked to a particular provider. In absence of such linked data, I suggest to delete “the study explore the link between PNC component coverage for specific service items and PNC provider quality of midwives”. 2. The limitations of the study may be placed at the end of the paper (may be as a second paragraph or second last paragraph of in the discussion section); it does not fit in the abstract. 3. In the method section, specify what was the unit for panel/follow-up - villages or households or women or midwives? It is mentioned “attrition of 4%”. But attrition of what? Specify if the attrition is of villages or households or women or midwives? 4. Although authors have given reference of design details, it will good to specify in one line how many households/women and midwives were interviewed at baseline (2007) and at endline (2009), or is it the same as mentioned in page 5 (line 16-19) 5. Table 1 is confusing. What are the four columns corresponding to treatment and control? At some places, I guess ‘n’ values are given corresponding to “Missing”. All these need to be specified in the column heading.

REVIEWER	Sarika Chaturvedi Savitribai Phule Pune University, Pune, India
REVIEW RETURNED	27-Jun-2017

GENERAL COMMENTS	The authors have revised the manuscript satisfactorily and is now suitable for publication.

REVIEWER	Ngianga II Kandala University of Portsmouth, UK None
REVIEW RETURNED	09-Jun-2017

GENERAL COMMENTS	I am satisfy with the Authors responses and the paper reads well. Good job.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 5
Reviewer Name: Ngianga II Kandala
Institution and Country: University of Portsmouth, UK
Please state any competing interests: None

Please leave your comments for the authors below

COMMENT: I am satisfy with the Authors responses and the paper reads well. Good job.

RESPONSE: Thank you.

Reviewer: 1
Reviewer Name: David ZOMBRE
Institution and Country: University of Montreal School of Public Health, Canada
Please state any competing interests: None declared

Please leave your comments for the authors below

Comment: The authors made very useful and appropriate modifications to the previous version. I have no additional suggestions on the manuscript at this time.

Comment: Thank you.

Reviewer: 2
Reviewer Name: Tara Patricia Cookson
Institution and Country: University of British Columbia, Canada
Please state any competing interests: None declared

Please leave your comments for the authors below

The authors made significant improvements to the paper and I recommend it for publication contingent upon a few very minor changes, listed below.

Comment: Pg 5 Line 15 grammatical error: "we includes" --> "we include"

Response: Thank you for pointing this out. We have made the change.

Comment: Pg 12 line 10 grammatical error: "This low proportion is may be due" --> "This low proportion may be due"

Response:: Thank you for pointing this out. We have made the change.

Comment: Pg 12 lines 9 - 11: You suggest that only 20% of women had pelvic exams perhaps due to "the infrastructure of the healthcare facility and cultural norms". This seems insufficient, and while your study could not delve into the other material and social determinants of healthcare quality, it is important to at least consider and preferably point to research that considers the reasons why women receive inadequate sexual and reproductive healthcare. "Infrastructure" is too broad to pin women not receiving important pelvic exams. If you can, point to research that does explore this within the Philippine context.

Response: We inferred ‘infrastructure’ based on the possibility of privacy concerns at some health care facilities. Based on the facility survey from another data set, the Indonesian Family Life Survey (reference added). We have changed the sentence to:
“This low proportion may be due to the possibility of limited examination rooms at healthcare facilities (since only 54% of facilities have a separate maternal and child health or family planning examination room) and cultural norms on reproductive health.”
We add a reference on cultural norms and reproductive health that does not specifically discuss Indonesia, but does use Indonesia as an example:
Burki, T., 2010. Culture: Cancer and cultural differences. Lancet Oncology, 11(12), p.1125.

Comment: Pg 12 lines 16-20: You suggest that women not receiving the full schedule of iron pills is likely due to poor compliance with PNC visits and iron supplementation. How do you draw this correlation? How do you know it was not because of stock-outs? And how do you know that following this compliance with PNC visits became a requirement?

RESPONSE: We have changed the sentences to:
“A 30-day supply of iron-folic acid pills should be given to women as part of every PNC visit. Only 12% of women reported receiving at least 90 iron-folic acid pills during pregnancy, although about 80% of women received iron-folic acid pills at least once during pregnancy. This large discrepancy suggests women received iron supplementation at least once during their PNC visit, but women may show poor compliance to PNC visits, causing them to not receive the iron supplementation, or women do not receive iron supplementation during their PNC visit due to providers’ omission or insufficient stocks. To address both PNC visits and iron supplementation, compliance with PNC visit guidelines became part of the CCT programme’s requirements.”

Comment: Pg 15 lines 50-53: I am not sure what you mean in these two sentences regarding price dimension versus quality dimension. In the following sentence you (correctly) identify that when incentives are only provided to service users and not service providers, there may not be a motivation to improve service quality. But does this contradict what you say before? This paragraph needs some clarification work.

Rsponse: We have changed the sentence to:
“When incentives are only provided to patients, we find improved health-seeking behaviour, but not necessarily improved health outcomes. In this setting, healthcare providers have no incentive to improve the quality of service provided, and this may partly explain the limited health improvements.”

Reviewer: 3
Reviewer Name: Dr. Avishek Hazra
Institution and Country: Population Council, India
Please state any competing interests: 'None declared'

Please leave your comments for the authors below

The article now reads much better than the previous version. However, below are few suggestions that the authors might consider for further improvement.

1. I have strong reservation in the statement that “the study is exploring the link between PNC component coverage for specific service items and PNC provider quality of midwives” (as mentioned in line 40-41 in page 3). The study has independently shown the effect of the program on PNC utilization and provider’s quality using two separate datasets. Nowhere in the link between the two has been established through data. Exploring the link is only possible if individual women data could be linked/merged/appended to corresponding provider data, and then assigning average score for each outcome to all individual women who are linked to a particular provider. In absence of such linked data, I suggest to delete “the study explore the link between PNC component coverage for specific service items and PNC provider quality of midwives”.

RESPONSE: We have changed the sentence to “This study extends earlier reports by exploring PNC component coverage for specific service items and PNC provider quality of midwives.”

2. The limitations of the study may be placed at the end of the paper (may be as a second paragraph or second last paragraph of in the discussion section); it does not fit in the abstract.

RESPONSE: The limitations have been simplified in the abstract: “Measurement error limits the interpretation of the study since women with older children might not accurately recall the services received during pregnancy.”.
Further detail is included in the data section:
“Recall bias and measurement error may have influenced data quality, but the relatively short time window of 24 months would tend to limit overall bias.”

3. In the method section, specify what was the unit for panel/follow-up - villages or households or women or midwives? It is mentioned “attrition of 4%”. But attrition of what? Specify if the attrition is of villages or households or women or midwives?

RESPONSE: We have changed the sentence to: ”The follow-up was conducted between October and December 2009, attrition at the household level was 4%.”

4. Although authors have given reference of design details, it will good to specify in one line how many households/women and midwives were interviewed at baseline (2007) and at endline (2009), or is it the same as mentioned in page 5 (line 16-19)

RESPONSE: We have included the following:
”The baseline included 4,700 pregnancies and deliveries between June 2005 and August 2007. The follow-up included 2,168 pregnancies and deliveries between October 2007 and December 2009. “
“There were 1,396 observations from midwives in public practice and 1,269 observations from private practice.”

5. Table 1 is confusing. What are the four columns corresponding to treatment and control? At some places, I guess ‘n’ values are given corresponding to “Missing”. All these need to be specified in the column heading.

RESPONSE: We include missing observations for covariates and outcomes in table 1.

Reviewer: 4
Reviewer Name: Sarika Chaturvedi
Institution and Country: Savitribai Phule Pune University, Pune, India
Please state any competing interests: None declared

Please leave your comments for the authors below

The authors have revised the manuscript satisfactorily and is now suitable for publication.
RESPONSE: Thank you.