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

TITLE (PROVISIONAL)	Does oral polio vaccine have non-specific effects on all-cause mortality? Natural experiments within a randomised controlled trial of early measles vaccine
AUTHORS	Aaby, Peter; Andersen, Andreas; Martins, Cesario; Fisker, Ane; Rodrigues, Amabelia; Whittle, Hilton; Benn, Christine

VERSION 1 - REVIEW

REVIEWER	Andrea Benedetti McGill University Canada
REVIEW RETURNED	29-Apr-2016

GENERAL COMMENTS	This paper aims to investigate non specific effects of oral polio vaccination via two natural experiments and an RCT. This paper is poorly written, poorly organized and consequently very hard to understand. 1. Figure 2 is separated from its header. I find it confusing. please try to make it more clear. 2. Similarly, the description of the natural experiments is confusing. Clarify. 3. Please improve the grammar of the patient involvement paragraph. 4. The Statistical Methods section includes results that belong in the Results section and statements that belong in the Discussion. Please fix. 5. The Cox PH models should be adjusted for potential confounders - but be careful, given the small numbers of deaths some models will not accomodate adjusting for confounders. 6. Before presenting results pertaining to the hypotheses of interest, first please present the characteristics of the study population, according to exposure status for the two natural experiments. 7. For the p-values for interaction - please present exact pvalues for potential interactions, not just the ones that were statistically significant or borderline significant. Also - please describe how these were estimated. 8.Table 1: Given the small number of deaths, it would be more interesting to include timing of administration of OPV0 as a continuous covariate rather than presenting stratum specific estimates. 9. Table 2: It is not clear to me what the interaction p value is measuring. 10. Calling these analyses "intention to treat" seems misleading since these are no longer randomized comparisons. 11. Please address as a limitation the small number of cases in

	some strata. 12. It is not clear if there was missing data/dropouts, and what strategy was used to deal with it. 13. It is unclear what information is available to be used as potential confounders.

REVIEWER	Anne Mobley Butler University of North Carolina, Chapel Hill, NC, USA
REVIEW RETURNED	20-May-2016

GENERAL COMMENTS	Recommendation: Revise/resubmit with major revisions required In the manuscript entitled, Does oral polio vaccine have non-specific effects on all-cause mortality? Natural experiments within a randomized controlled trial of early measles vaccine, Aaby et al. sought to provide important evidence on possible non-specific effects (NSE) of OPV. The study provides estimates of the effect of oral polio vaccine (OPV) administration on the mortality rate ratio comparing two-dose (administered at 4.5 and 9 months) vs. one-dose of measles vaccine (MV) (administered at 9 months only). This manuscript has important strengths. • First, the study design is unique and novel. The authors took advantage of two natural experiments that occurred during a RCT to answer an important research question regarding non-specific effects of the trivalent OPV. • The authors have identified important information and citations to tell a good story. • The results stratified by timing of OPV is compelling, though 95% CIs are wide. However, the manuscript has major weaknesses: • Overall, the manuscript includes important pieces of information and is well-cited. But the bottom-line is that the writing needs major improvement. The story gets lost behind confusing logic and unconventional scientific writing. For example, the justification in the introduction needs improvement. Authors need to clarify whether previous analysis that censored measles deaths was appropriate statistical analysis. The language is ambivalent. Also, one wonders whether the authors meant that the lack of indication that prevention of BCG explains the (ENTIRE?) beneficial effect. The science is well-thought out but needs attention. • Results section should begin with brief description of study population. • How well-balanced were the two groups with respect to covariates? Randomization does not always result in perfectly balanced groups – thus, this information needs to be reported. If balance is not achieved, adjustment will be necessary in the analyses. • The authors failed to address whether the nonspecific effects differ by gender, which is a frequently observed in studies of nonspecific effects of vaccines. Was gender data available? Is it possible to perform this analysis? At the very least, the subject warrants a comment in the discussion. • The non-specific effects of vaccines can be boosted or diminished when other immunomodulating health interventions (e.g., other vaccines, or vitamins) are provided. The figure demonstrates some

	thought in this regards, but was this incorporated into the analyses? • The authors should consider mentioning in the discussion section the international replacement (which began April 2016) of trivalent OPV with bivalent OPV, as new questions will arise regarding the bivalent OPV: (http://www.who.int/immunization/diseases/poliomyelitis/endgame_objective2/oral_polio_vaccine/en/) • Box 1 provides important information – but this is presented in a way that is difficult for the reader to interpret. Either, the table needs to be reorganized; or deleted and transfer key information ought to intro and discussion sections. • Table 1: split rate (deaths/person-days) and N into two different columns. Is that a 95% CI in the MRR column? Need to label. • Figure 1 clearly communicates the study design. However, the schematic seems to indicate that follow-up occurred AFTER not BEFORE or AT 36 months. This is misleading. • Figure 2 is unclear. It requires further description. Abbreviations need to be defined. The text is poor quality (particularly "periods with missing OPV0"). • The results would be drastically improved using data visualization. Table 1 • Authors should consider displaying results in a figure form – difficult to interpret all data in tables.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Andrea Benedetti
Institution and Country: McGill University, Canada Competing Interests: None declared.

This paper aims to investigate non specific effects of oral polio vaccination via two natural experiments and an RCT.

This paper is poorly written, poorly organized and consequently very hard to understand.

1. Figure 2 is separated from its header. I find it confusing. please try to make it more clear.

PA: We have revised the legend for figure 2 hopefully making it clearer.

2. Similarly, the description of the natural experiments is confusing. Clarify.

PA. We have tried to improve the description of the two natural experiments.

3. Please improve the grammar of the patient involvement paragraph.

PA: The paragraph has been rewritten.

4. The Statistical Methods section includes results that belong in the Results section and statements that belong in the Discussion. Please fix.

PA: It was not clear to me which results in the methods section the reviewer wanted to move to the result section and which comments to the discussion. The "results" in the method section are descriptions of the RCT, and serve as background information; they are not results of the present analysis.

5. The Cox PH models should be adjusted for potential confounders - but be careful, given the small numbers of deaths some models will not accomodate adjusting for confounders.

PA: The original RCT description (ref 8) had a table with potential confounders. Since this was an RCT and the potential confounders did not differ between randomization groups, the original analysis (ref 8) only stratified for the three health centre areas where enrolment took place. We have followed the same strategy in the present analysis. Since a main point of the present description is to see

whether campaigns change the results of RCTs it would be wrong if we started controlling for another set of background factors. This has now been explained in the methods section.

6. Before presenting results pertaining to the hypotheses of interest, first please present the characteristics of the study population, according to exposure status for the two natural experiments.
PA: This has now been done.

7. For the p-values for interaction - please present exact pvalues for potential interactions, not just the ones that were statistically significant or borderline significant. Also - please describe how these were estimated.
PA: It was said in the methods section: "To test for no interaction we compared the effect of two-dose versus one-dose MV in strata of the suspected effect modifier using Wald statistics." We present interaction tests for all comparisons made in the text.

8.Table 1: Given the small number of deaths, it would be more interesting to include timing of administration of OPV0 as a continuous covariate rather than presenting stratum specific estimates.

PA: This information has been added in the result section: "If age-at-OPV0 was analysed as a continuous variable, the MRR (two-dose/one-dose MV) estimate was reduced by 4% (1-6%) for each additional day of age at the time of vaccination."

9. Table 2: It is not clear to me what the interaction p value is measuring.
PA: The interaction test examined whether the MRRs for 2-doses/1-dose of MV were the same for children who had received OPV before enrolment and those who had not. This is now mentioned in the footnote.

10. Calling these analyses "intention to treat" seems misleading since these are no longer randomized comparisons.
PA: The concept Intention to treat has been removed.

11. Please address as a limitation the small number of cases in some strata.
PA: This is now mentioned in the discussion.

12. It is not clear if there was missing data/dropouts, and what strategy was used to deal with it.
PA: Follow-up/dropouts was described in the original trial description (BMJ 2010, ref 8). Drop-out has been handled in the same way as in the original analysis, i.e. children moving or dying have been censored at the data of movement or death. We assumed that to be implicit but this is now mentioned specifically in the methods section.

13. It is unclear what information is available to be used as potential confounders.
PA: The original RCT description (ref 8) had a table 1 with potential confounders. Since this was an RCT and the potential confounders did not differ between randomization groups, the original analysis (ref 8) only stratified for the three health centre areas where enrolment took place. We have followed the same strategy in the present analysis. Since a main point of the present description is to see whether campaigns change the results of RCTs it would be wrong if we started controlling for another set of background factors. This has now been explained in the methods section.

Reviewer: 2
Reviewer Name: Anne Mobley Butler
Institution and Country: University of North Carolina, Chapel Hill, NC, USA Competing Interests: None declared

Recommendation: Revise/resubmit with major revisions required

In the manuscript entitled, Does oral polio vaccine have non-specific effects on all-cause mortality? Natural experiments within a randomized controlled trial of early measles vaccine, Aaby et al. sought to provide important evidence on possible non-specific effects (NSE) of OPV. The study provides estimates of the effect of oral polio vaccine (OPV) administration on the mortality rate ratio comparing

two-dose (administered at 4.5 and 9 months) vs. one-dose of measles vaccine (MV) (administered at 9 months only).

This manuscript has important strengths.

- First, the study design is unique and novel. The authors took advantage of two natural experiments that occurred during a RCT to answer an important research question regarding non-specific effects of the trivalent OPV.
- The authors have identified important information and citations to tell a good story.
- The results stratified by timing of OPV is compelling, though 95% CIs are wide.

PA: Thanks

However, the manuscript has major weaknesses:

- Overall, the manuscript includes important pieces of information and is well-cited. But the bottom-line is that the writing needs major improvement. The story gets lost behind confusing logic and unconventional scientific writing. For example, the justification in the introduction needs improvement. Authors need to clarify whether previous analysis that censored measles deaths was appropriate statistical analysis. The language is ambivalent. Also, one wonders whether the authors meant that the lack of indication that prevention of BCG explains the (ENTIRE?) beneficial effect. The science is well-thought out but needs attention.

PA: We have tried to improve the introduction. It is correct that the previous analyses censoring for measles death or measles infection were appropriate survival analyses in which follow-up was censored at the time of infection. This has now been mentioned and references have been added.

- Results section should begin with brief description of study population.

PA: This has been added

- How well-balanced were the two groups with respect to covariates? Randomization does not always result in perfectly balanced groups – thus, this information needs to be reported. If balance is not achieved, adjustment will be necessary in the analyses.

PA: Balance was achieved - see Table 1 in ref 8. This is now justified with reference to the original trial paper.

- The authors failed to address whether the nonspecific effects differ by gender, which is a frequently observed in studies of nonspecific effects of vaccines. Was gender data available? Is it possible to perform this analysis? At the very least, the subject warrants a comment in the discussion.

PA: We do have the data by sex. We have now mentioned in the results section that the effect of two-doses of MV was similar for boys and girls among children who had not received campaign-OPV and likewise for the children who had received campaign-OPV before enrolment in the trial. However, among the children who received OPV after enrolment, the effect of two-doses was significantly better for girls than for boys (p=0.05). This has also been mentioned in the revised manuscript. Campaign-OPV after enrolment was usually given with VAS since most children had reached 6 months of age. The sex-difference in the effect of two-dose MV was particularly strong among children who had received campaign-OPV with VAS, the MRR (two-dose/one-dose MV) being 0.27 (0.09-0.77) for girls and 1.46 (0.69-3.09) for boys (p=0.01). This is a credible interaction since we have previously shown in a randomized trial that VAS administered with MV was associated with significantly lower mortality for girls whereas for boys it was associated with increased *mortality (ref 16). At the moment we have not mentioned this in the manuscript as we thought it might complicate the presentation too much with yet another interaction. However, if the editors think this is important we can of course add this information.*

- The non-specific effects of vaccines can be boosted or diminished when other immunomodulating health interventions (e.g., other vaccines, or vitamins) are provided. The figure demonstrates some thought in this regards, but was this incorporated into the analyses?

PA: A very good point. Yes. The analyses dealt with the effects of the OPV campaigns occurring during the conduct of the two-dose trial. There was also a measles vaccination campaign for children aged 6 months to 15 years of age in 2006. However, the children enrolled in the two-dose trial were exempted from participation in this campaign, so it should not have had effect on the present analysis. The measles vaccination campaign is now mentioned in the methods section.

The other campaigns during the conduct of this trial were VAS campaign in November 2003, and VAS and mebendazole campaigns in May and December 2006, July and December 2007, July 2008, January and July of 2009. These campaigns did not affect the mortality rate within the MV trial; the mortality rate after VAS-only campaign versus before VAS campaigns was 0.92 (0.63-1.35). This is now mentioned in the methods section.

- The authors should consider mentioning in the discussion section the international replacement (which began April 2016) of trivalent OPV with bivalent OPV, as new questions will arise regarding the bivalent OPV: (http://www.who.int/immunization/diseases/poliomyelitis/endgame_objective2/oral_polio_vaccine/en/)

PA: In analyses of subsequent OPV campaigns we found that there were similar effects of monovalent, bivalent and trivalent OPV. This is now mentioned in the discussion.

- Box 1 provides important information – but this is presented in a way that is difficult for the reader to interpret. Either, the table needs to be reorganized; or deleted and transfer key information ought to intro and discussion sections.

PA: Box 1 has now been reorganized to a hopefully more clear presentation.

- Table 1: split rate (deaths/person-days) and N into two different columns. Is that a 95% CI in the MRR column? Need to label.

PA: This has now been done. Yes it is 95% CI.

- Figure 1 clearly communicates the study design. However, the schematic seems to indicate that follow-up occurred AFTER not BEFORE or AT 36 months. This is misleading.

PA: We have now modified the figure to make it more clear.

- Figure 2 is unclear. It requires further description. Abbreviations need to be defined. The text is poor quality (particularly "periods with missing OPV0").

PA: We have provided further explanation in the legend to Figure 2 – hopefully making it easier to understand.

- The results would be drastically improved using data visualization. Table 1
- Authors should consider displaying results in a figure form – difficult to interpret all data in tables.

PA: We have presented the data in Table 1 also as a graph: Figure 3. – We hope this is what the reviewer wants. *I am a number person rather than a visualization person so I would prefer to keep Table 1 in the text but if the editor thinks that is better Table 1 can be transferred to a Supplementary Table.*

VERSION 2 - REVIEW

REVIEWER	Andrea Benedetti McGill University, Canada
REVIEW RETURNED	11-Jul-2016

GENERAL COMMENTS	The authors have addressed several of my original comments, and overall the paper is more clearly described and easier to understand now. However, there are still a number of concerns that have not been adequately addressed. 1. In my first review, I asked that the study population be compared with respect to relevant confounders according to the natural experiments. In fact, what seems crucial is that the confounders are balanced within strata. Please present tables addressing this. If confounders were unbalanced within strata, this could affect the results and these should be adjusted for. Moreover, the description of the population is an important aspect of this paper, and should be presented as part of this paper, not via referencing another paper. 2.For the p values for interaction – it is still unclear to me. For

	example, Table 1 presents a pvalue =0.03, with asterisks on two categories. I am not sure what is going on here – is this the overall p value for the 4 strata – or for the difference between the 2 starred strata? What about p values for the other strata or were these not assessed. How these p values are described and presented still needs work. 3.Was an intention to treat analysis followed? (I originally asked that this wording be removed – however, at the beginning of the results section, they present per protocol results, which should usually be adjusted for confounders. All in all, I am unsure what was done. Please specify.)

REVIEWER	Anne Butler University of North Carolina USA
REVIEW RETURNED	01-Aug-2016

GENERAL COMMENTS	The authors failed to address many of the previous criticisms. Also, the writing is difficult to understand.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Andrea Benedetti
Institution and Country: McGill University, Canada Competing Interests: None declared.

The authors have addressed several of my original comments, and overall the paper is more clearly described and easier to understand now.

However, there are still a number of concerns that have not been adequately addressed.

1. In my first review, I asked that the study population be compared with respect to relevant confounders according to the natural experiments. In fact, what seems crucial is that the confounders are balanced within strata. Please present tables addressing this. If confounders were unbalanced within strata, this could affect the results and these should be adjusted for. Moreover, the description of the population is an important aspect of this paper, and should be presented as part of this paper, not via referencing another paper.

PA: We have now compared the potential confounders in the two analyses: a) for recipients of OPV0 versus non-recipients of OPV0 – corresponding to Table 1; and b) for those who received and did not receive campaign-OPV before enrolment – corresponding to Table 2. This information has been presented in the text and in Supplementary Tables 1 and 2. We also present analyses where we have adjusted for the background factors which are likely to have differed significantly prior to allocation to the two comparison groups.
A) In the analysis related to reception of OPV0 there were small but statistically significant differences in distribution of sex, mother's mid-upper-arm-circumference (MUAC), and in fever, diarrhoea, child's MUAC and height at enrolment. Only sex and mother's MUAC are likely to have been different at the time of allocation and therefore we have only adjusted for those in the analysis presented in the Supplementary table. Instead of the MRRs of 0.45 (0.29-0.71), 0.78 (0.36-1.70), 1.10 (0.47-2.59) and 3.63 (0.87-15.2) presented in Table 1, we now get 0.47 (0.30-0.74), 0.79 (0.36-1.72), 1.20 (0.51-2.87) and 3.43 (0.82-14.38) in the adjusted analysis (Supplementary Table 3). The differences in fever, diarrhoea, child's MUAC and height at enrolment could be a result of the group allocation and should

therefore not be treated as a confounder but even if we do adjust for all those differences in the analyses the results are essentially the same: we get the MRRs 0.48 (0.31-0.77), 0.80 (0.37-1.76), 1.20 (0.50-2.87), and 3.47 (0.83-14.56). These latter results have not been presented in the paper. B) In relation to reception of campaign-OPV before enrolment there were small but statistically significant differences in distribution of district and age of child (which are already adjusted for in the analysis), number of persons sleeping per room, mother's MUAC, and in fever, diarrhoea, and respiratory frequency at enrolment. Only number of persons sleeping per room and mother's MUAC are likely to have been different at the time of allocation and we have therefore only adjusted for those in the analysis presented in the Supplementary table: Instead of the MRR of 0.60 (0.42-0.85) and 1.16 (0.64-2.13) (interaction test, p=0.06) we now get: 0.61 (0.43-0.87) and 1.19 (0.64-2.22) (interaction test, p=0.07) (Supplementary Table 4). The differences in fever, diarrhoea, and respiratory frequency at enrolment could be a result of the group allocation and should therefore not be treated as a confounder but even if we do adjust for all those differences in the analyses the results are essentially the same: we get the MRRs 0.62 (0.43-0.88) and 1.32 (0.71-2.43) (interaction test, p=0.04). These latter results have not been presented in the paper.

2. For the p values for interaction – it is still unclear to me. For example, Table 1 presents a pvalue =0.03, with asterisks on two categories. I am not sure what is going on here – is this the overall p value for the 4 strata – or for the difference between the 2 starred strata? What about p values for the other strata or were these not assessed. How these p values are described and presented still needs work.

PA: The p-value=0.03 was based on a comparison of the two strata with an asterisk. We have now removed this test as it apparently creates confusion and merely emphasized that there is a significant trend with age at OPV0 vaccination. The test used in Tables 1 and 2 have been explained in greater details in notes to Tables 1 and 2.

3.Was an intention to treat analysis followed? (I originally asked that this wording be removed – however, at the beginning of the results section, they present per protocol results, which should usually be adjusted for confounders. All in all, I am unsure what was done. Please specify.)

PA: The result presented in the beginning of the result section are the PP results for children who did receive two doses of MV comparing early two-dose with controls as presented in the original paper (ref 3). This has been further explained in the beginning of the result section. Our key interest is not to adjust this estimate for possible confounding but to examine whether the estimate changes due to the OPV-campaigns. However, we have also made the analyses corresponding to tables 1 and 2 in relation to the ITT data set and placed these results in Supplementary Tables 5 and 6. It made no difference. This has now been noted in the result section.

Reviewer: 2
Reviewer Name: Anne Butler
Institution and Country: University of North Carolina, USA Competing Interests: None declared

The authors failed to address many of the previous criticisms. Also, the writing is difficult to understand.

VERSION 3 – REVIEW

REVIEWER	Andrea Benedetti McGill University
REVIEW RETURNED	03-Nov-2016

GENERAL COMMENTS	The presentation is clearer and I am more satisfied with the additional results presented. My review has mainly addressed how the data were analyzed. This article should also be reviewed by a subject matter expert. A few additional questions: 1. The tables are not very clear. For Table 1, the MRR column should state that there is a 95% confidence interval in the brackets after the MRR. Add to the footnotes that these estimated came from a Cox PH model. What are the values presented in the second and third column - for example, for all children in the early two dose group, the Mortality rate is 1.23 with 58 deaths in 1722488 person years among 2129 children. Where does the 1.23 come from - why isnt it 58/1722488? Clarify. Also, the age-trend did not test whether the effect was equal in the 4 groups, but rather whether the trend increased in a linear way. Please make the same clarifications for Table 2 and 3, and the tables presented in the supplemental materials.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Andrea Benedetti
Institution and Country: McGill University, Canada Competing Interests: None declared.

The presentation is clearer and I am more satisfied with the additional results presented. My review has mainly addressed how the data were analyzed. This article should also be reviewed by a subject matter expert.

A few additional questions:

1. The tables are not very clear.

For Table 1, the MRR column should state that there is a 95% confidence interval in the brackets after the MRR.
Add to the footnotes that these estimated came from a Cox PH model.
PA: this has been revised as suggested in all tables.

What are the values presented in the second and third column - for example, for all children in the early two dose group, the Mortality rate is 1.23 with 58 deaths in 1722488 person years among 2129 children. Where does the 1.23 come from - why isnt it 58/1722488? Clarify.
PA: The reviewer may have misread the tables. The headings said “deaths/person-days” not “person-years”. Hence, 58/1722488*365.25=0.0123 (or 1.23). To prevent the misunderstanding we added in the footnotes: “to be precise we have reported person-days and not person-years”.

Also, the age-trend did not test whether the effect was equal in the 4 groups, but rather whether the trend increased in a linear way.

PA: The age-trend tested for a significant linear increase in the effect across the 4 groups.This has been noted in the footnotes in the relevant tables.

Please make the same clarifications for Table 2 and 3, and the tables presented in the supplemental materials.
PA: This has been done as suggested