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)	Cluster randomised controlled trial to examine medical mask use as source control for people with respiratory illness
AUTHORS	MacIntyre, Raina; Zhang, Yi; Chughtai, Abrar; Seale, Holly; Zhang, Daitao; Chu, Yanhui; Zhang, Haiyan; Rahman, Bayzidur; Wang, Quanyi

VERSION 1 - REVIEW

REVIEWER	Donald Milton University of Maryland School of Public Health USA
REVIEW RETURNED	18-May-2016

GENERAL COMMENTS	Thinking out side of the box (in this case a respirator) about alternatives to personal protective devices (PPE) is an important step for influenza infection control. When vaccines are not available or are mismatched, exposure control becomes critical. The authors rightly point out the difficulties with maintaining compliance among medical workers using PPE. And, the larger issue is what to do about protecting other patients in the medical setting and in the larger community. Thus, this study addresses an important issue by testing source control through the use of masks by index cases. The major limitations of the study are two: lack of power and weak follow-up design. The design expectation of 10-20% secondary attack rate (SAR) was not unreasonable based on previous RCTs in Hong Kong and Bangkok[1,2]. The expected effect, however, was probably too large. Surgical masks have been shown to reduce influenza aerosol concentrations by a factor of a little over 2, with most of the effect limited to large droplet spray.[3] Because index cases were not expected to wear the masks all of the time, the shedding would be expected to be reduced by less than the 50% best case and this might be expected to result in an attenuated effect. The bigger problem for this study was that the observed SAR was 1-2% (only 10% of that anticipated). It is not clear why the observed rate was so low. A possible reason is that unlike the Hong Kong and Bangkok studies, the primary outcome relied on symptom reports rather than collecting swabs to identify all infections – symptomatic or not. The methods are somewhat vaguely described – so it is not clear if all households were visited once or more during the intervention and how the diary cards were returned or where the symptomatic cases were swabbed (at home or in the fever clinic or elsewhere). The authors do not attempt to explain why the SAR was so low and out of line with other studies. In the end, only one of the three endpoint measures in a sort of per protocol analysis showed a significant effect.

	This is not the first RCT of masks as source control as claimed in the discussion; Two studies where masks were used as both source control and PPE[2,4] are cited above, and MacIntyre goes on to describe a previous one done by some of the same authors[5] as well as discuss one done in France[6] that specifically addressed source control. A revised manuscript would need to give a clearer description of the follow-up procedures and discuss why the observed SAR was so low. However, given the severely limited power as a result of the low SAR, and the weak finding for mask use per and regardless of protocol, it is sadly to say that this study cannot make an important contribution to what we know about influenza prevention. Reference Cited: 1 Cowling BJ, Chan KH, Fang VJ, et al. Facemasks and Hand Hygiene to Prevent Influenza Transmission in Households: A Randomized Trial. Ann Intern Med Published Online First: 2009.http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve& 2 Simmerman JM, Suntarattiwong P, Levy J, et al. Findings from a household randomized controlled trial of hand washing and face masks to reduce influenza transmission in Bangkok, Thailand. Influenza Other Respir Viruses 2011;5:256–67. doi:10.1111/j.1750-2659.2011.00205.x 3 Milton DK, Fabian MP, Cowling BJ, et al. Influenza virus aerosols in human exhaled breath: particle size, culturability, and effect of surgical masks. PLoS Pathog 2013;9:e1003205. doi:10.1371/journal.ppat.1003205 4 Cowling BJ, Chan KH, Fang VJ, et al. Facemasks and Hand Hygiene to Prevent Influenza Transmission in Households: A Randomized Trial. Ann Intern Med Published Online First: 2009.http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve& 5 MacIntyre CR, Chughtai AA. Facemasks for the prevention of infection in healthcare and community settings. BMJ 2015;350:h694. doi:10.1136/bmj.h694 6 Canini L, Andréoletti L, Ferrari P, et al. Surgical mask to prevent influenza transmission in households: a cluster randomized trial. PloS One 2010;5:e13998. doi:10.1371/journal.pone.0013998

REVIEWER	Gary Garber Public health Ontario Canada
REVIEW RETURNED	19-May-2016

GENERAL COMMENTS	The authors performed a cluster-randomised trial to examine the effect of masking of infected participants on the transmission of respiratory infections (and colonization for bacteria) to household members. This paper is a needed addition to the lacking literature regarding source control using medical masks.

	Overall, the study was described clearly and the data appears to be coherent with authors' conclusions in a balanced manner with an impressive 0% loss to follow-up. Understandably, attack rates are generally low in these types of studies, so it is not surprising to see a lack of statistical significance. However, some items for consideration have been noted below, particularly for elements of the methods that are not clear. The paper needs more clarity of what constituted the post hoc analysis particularly as CLI reduction was sig. The discussion should discuss the high hand hygiene rates in general in this cohort and the higher rates among the "control " arm which likely had an imp[act on the number of event that were being counted Comments: <u>General</u> Is there any confounding due to treatment received by patients to clinics prior to and following recruitment that would reduce the severity of symptoms? No testing of study subjects for infection during recruitment: ILI used. While a small limitation, there would be no way to confirm that infections in households are the same infectious agent. It is understood that this is often not completed due to intensive resource allocations required. Please add a line to comment on "laboratory-confirmed bacterial colonization" and how that outcome is not included in your paper, but was listed as an outcome for your trial Protocol. <u>Specific</u> Page 2, Abstract: Identify design as "Cluster-randomised controlled trial". Page 5 line 1-2: suggest to specify in a community setting, "… whether medical mask use by people *in a community setting* with ILI…" Generalization of data from this study to health care facilities would need significant consideration of variables that differ between settings, which impact transmission (e.g., cues for compliance, proximity of other susceptible individuals [patients, HCW, visitors],

	ventilation dynamics of the room, adherence to other contact precautions such as hand hygiene). Nevertheless, that discussion would be beyond the scope of this paper. Page 6, Line 2: 225 index cases included (also mentions 225 in abstract), but 245 are noted later in the paragraph (123+122) and in Figure 1. Please correct. Page 6, Line 3-5: Please provide greater detail on methods for sequence generation. Page 6, line 14: In consideration of the cultural context in Beijing, is there any risk of contact transmission during meal times wherein index cases may have shared use of utensils or are likely to have contaminated shared food sources (e.g., 'buffet-style' dining?) that is necessary to note? Page 6, line 15: Was proper hand hygiene technique shown? Page 6-7, Outcome measures: Would you clarify who performed daily measurements? Were index cases required to report daily whether or not household contacts had symptoms? Or was no notice from index cases assumed to be indication of no household contact having symptoms? If there are no data being collected daily except when a notification occurs, then is this really "Primary endpoints measured in household contacts on a daily basis"? (It appears that some answers appear in the "Data collection and follow-up" section, but clarity sooner would be better.) Page 6-7, outcome measures: CRI is a very broad definition that would be inclusive of ILI. CRI was used as an outcome – one that was significant in post-hoc analysis; yet, CRI was not used as an eligibility criterion for selecting index cases. Would you add rationale for the use of CRI as an outcome, but not for eligibility? To my knowledge, CRI is not a standard outcome other than having been used in previous publications by your colleagues. Because ILI is used for selection, I feel that it would be more balanced to include in the abstract, in addition to CRI post-hoc analysis data that is already mentioned, the post-hoc analysis data from the results section related to ILI comparison of mask group vs no mask group. Page 8, Data collection and follow-up: Regarding accuracy of CRI

	and ILI outcome measurements: When were diary cards collected? Were symptoms of household members assessed by index cases, then confirmed by a (medically) trained project member during swabs for laboratory detection? Page 10, Line 10-13: Regarding the validity for mask and no-mask group assignment for the post-hoc analysis: Were compliance measures for mask use or lack of mask use measured in the control group? How was the mask use identified (by diary or by interview)? How many times must a mask have been used to be eligible for inclusion in the “mask group” of the post-hoc analysis? Page 11, Line 20-21: Please add the number of index cases that did NOT use a mask during the trial period that were from the mask arm.

REVIEWER	Lisa Brosseau University of Illinois at Chicago United States
REVIEW RETURNED	20-May-2016

GENERAL COMMENTS	General Comments This paper is a community study of the impact of medical masks worn by ill household members on illness in other household members. The inclusion of information in the Introduction and Discussion about the role of medical masks in healthcare settings is confusing and not clearly tied to the goals or context of this study. I recommend focusing only on research and data related to the wearing of medical masks by sick people in community settings (i.e. homes), which are very different from healthcare settings. A brief mention of the original purpose of a medical mask – worn by healthcare workers to prevent infections in patients resulting from exposure to emissions from healthcare workers – is all that is needed. I am not convinced that healthcare settings have any relevance to home settings, because the nature of interactions between household members and the environmental conditions (e.g. air flow) are entirely different. Thus, the wearing of masks in healthcare settings – by healthcare workers or patients – is entirely unrelated to the wearing of masks in community settings. Abstract, Strengths & Limitations: • Which of these are strengths and which are limitations? • Pt 2: This point is confusing and poorly written. Please revise for clarity. Introduction • Lines 9-11: Medical masks are not worn in healthcare settings to protect healthcare workers in the U.S. That is the function of a respiratory protection device. The purpose of a surgical or medical mask is the protection of patients from healthcare worker emissions. • Paragraph 2: This paragraph is confusing – it mixes studies

	evaluating the effectiveness of masks worn by patients with effectiveness of masks worn by healthcare workers. I suggest separating into two different paragraphs focusing on these two different goals. • Lines 16-17: There have been at least 3 clinical trials of the effectiveness of medical masks on preventing surgical wound infections – one of these is cited (Tunevall). The others are: o Pippin DJ, Verderame RA, Weber KK. Efficacy of face masks in preventing inhalation of airborne contaminants. J Oral Maxillofac Surg 1987;45:319-23. o Davies KJ, Herbert AM, Westmoreland D, Bagg J. Seroepidemiological study of respiratory virus infections among dental surgeons. Br Dent J 1994;176:262-5. • Lines 43-48: Masks worn by whom and infection in whom? (see general note on Paragraph 2 above). Patient Involvement (Page 7) • This section is confusing, particularly the part related to publishing in an open journal. Discussion This section is generally well-written and organized, particularly where the discussion is focused on household and community trials. I think the two paragraphs on pages 13-14 and the final paragraph in this section that discuss mask use in healthcare settings should be minimized or deleted altogether. I think there is little to be gained from drawing comparisons between community studies of sick people wearing masks and healthcare settings with either patients or workers wearing masks. The settings and situations are too different.

REVIEWER	Mona Kanaan, PhD Senior Lecturer Programme Lead PGCert in Health Research and Statistics Department of Health Sciences University of York Heslington, York YO10 5DD, England
REVIEW RETURNED	18-Jul-2016

GENERAL COMMENTS	This manuscript reports the results of a randomised controlled trial (RCT) that examined the use of medical masks as source control for people with respiratory illness in a Chinese setting. The question is an interesting one. However, there are a number of issues that the authors might want to address. I list the major points then follow with the minor points. ITT (main analysis) Vs. the post-hoc analyses -In the abstract the results for the post-hoc analyses should be clearly labelled as such so that the reader is aware of potential biases. Similarly, in the main text the main and the post-hoc analyses are intertwined these should be clearly separated. - The post-hoc analyses are potentially biased as they break the randomisation. A CACE analysis could have been a better alternative [Connell AM. Employing Complier Average Causal Effect Analytic Methods to Examine Effects of Randomized Encouragement Trials. The American journal of drug and alcohol abuse. 2009;35(4):253-259. doi:10.1080/00952990903005882.] -It is not clear how time was measured based on the outcome measures section. As it is currently presented the mention of Cox

	regression and Kaplan Meier curves in the “Data analysis” section is questionable. Also, I wonder whether the one can afford a multivariable regression model given the very few events encountered. As a rule of thumb, one needs at least ten events per explanatory variable for a reliable estimate. Even when this is relaxed, at least five events were needed. [Vittinghoff E and McCulloch CE. Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression Am. J. Epidemiol. (2007) 165 (6): 710-718 first published online December 20, 2006 doi:10.1093/aje/kwk052]. This could explain the reported wide confidence intervals. - Page 10 Lines 13/15 this sentence needs revisiting as it is not clear why you would use information from a multiple regression type model to inform univariate analysis. -The use of person-day rates should be clearly stated in the analysis section. Also, were rates used in the sample size calculation or were these based on proportions? On a related note, the authors mention on Page 6 Line 43/45 that participants were informed that they could stop wearing a mask once their symptoms have resolved, how was this documented and how did they account for this when calculating person-day data. A Cluster RCT: - It should be clear in the title that this is a cluster randomised trial. - It is not clear whether the main analyses were made at the individual level or at the cluster (household) level. It should be clearly stated in this case how clustering was adjusted for in a similar fashion to how this was mentioned for the Cox regression. - The modified CONSORT for cluster randomised trial should have been used. - Reporting of ICCs is advised to aid sample size calculations for future studies. - Page 11 Line 11/15 a clarification whether these were 4 members from 4 different households or whether some were from the same household should be clarified. Ditto for the ensuing sentence Minor Points: - In the discussion, reasons for the study being under-powered are worthy of discussion in order to inform larger trials. Also it would be interesting to discuss the discrepancy between the observed proportions/rates of the main outcomes in this study and the assumed ones used in the sample size calculations. -Table 1 should only include baseline data. Post randomisation data should be presented in a separate table, such as “average hour mask wearing”. - P13 L15 replace “to examine to” with “to examine the”

VERSION 1 – AUTHOR RESPONSE

REVIEWER 1:

1. Why the observed SAR were low
Authors’ reply: Secondary attack rates (SAR) may be lower due to testing symptomatic cases only. We have added following sentence in the limitation section, “However the secondary attack rates

were much lower in this study which might be due to testing only symptomatic cases". Please see page 17 last para and 18 para 1.

2. It is not clear if all households were visited once or more during the intervention?
Authors' reply: All household members were visited at the start of the study only. Index cases were asked to report if any symptoms appeared in household members. Staff in the district CDC also contacted index cases via telephone every second day to check whether any household member developed symptoms.
We have modified the following para, "Symptoms in the family members were also recorded in the diary cards and index cases were asked to report any symptom. The index cases were asked to contact the study coordinator if any of the following symptoms appeared in household members: cough, nasal congestion, runny nose, sore throat, sneezes, chill, lethargy, loss of appetite, abdominal pain and muscle or joint aches. The study coordinator then assessed the household member and completed a follow-up survey. Samples obtained from all symptomatic cases. All index cases in the intervention and control arms were also asked to document compliance with mask use. Diary cards to record mask use were given to each index case, and they were asked to carry them during the day. Diary cards were returned to the investigators at the end of the study. Staff in the district CDC also contacted index cases via telephone on every alternate day to check whether any household member developed symptoms". Please see page 9 last para and page 10 first para.

3. How the diary cards were returned
Authors' reply: Diary cards were returned to the investigators at the end of the study. We have added following statement, "Diary cards were returned to the investigators at the end of the study" Please see page 10 first para.

4. Where the symptomatic cases were swabbed (at home or in the fever clinic or elsewhere)?
Authors' reply: Symptomatic cases were swabbed at the home by the trained investigators. We have added following statement, "Swabs were taken at the home by trained investigators". Please see page 10 para 2.

5. This is not the first RCT of masks as source control as claimed in the discussion.
Authors' reply: We have removed this statement (please see page 15, para 3) and also added the suggested references.

REVIEWER: 2

1. Is there any confounding due to treatment received by patients to clinics prior to and following recruitment that would reduce the severity of symptoms?
Authors' reply: We asked index cases about the history of medication (antibiotic, anti-tussive, anti-pyretic and other traditional medicine). The use of medication was not significantly different in two groups. We have modified statement to, "Some differences were noted between arms, but most characteristics, including medication use (data not shown), were generally similar between the two groups". Please see page 13 last para.

2. No testing of study subjects for infection during recruitment: ILI used. While a small limitation, there would be no way to confirm that infections in households are the same infectious agent. It is understood that this is often not completed due to intensive resource allocations required.
Authors' reply: We collected laboratory samples from index patients at the time of recruitment and from symptomatic household members during follow up. We have provided the following detail about the pathogens, "Two laboratory confirmed infections were identified among symptomatic household members – only one had the same infection (influenza H1N1) as the respective index case. Rhinovirus was isolated from other household member however no pathogen was isolated from

respective index case". Please see page 14 para 1.
We have also added following paragraph detailing the pathogens isolated in index patients, "Viruses were isolated from 60% (146/245) index cases. Influenza was the most common virus isolated from 115 (47%) cases - Influenza A - 100, Influenza B - 11 and Influenza A&B - 4. Other viruses isolated from index cases were, rhinovirus (14), NL63 (12) and C229E (7). More than one virus was isolated in 48 (20%) index cases, including 17 coinfections with influenza". Please see page 13 last para.

3. Please add a line to comment on "laboratory-confirmed bacterial colonization" and how that outcome is not included in your paper, but was listed as an outcome for your trial Protocol.
Authors' reply: Laboratory-confirmed bacterial colonization was included as outcome in the initial protocols however we did not test for bacteria due to limited budget resource constraints.

4. Page 2, Abstract: Identify design as "Cluster-randomised controlled trial".
Authors' reply: We have added this. Please see page 2.

5. Page 5 line 1-2: suggest to specify in a community setting, "… whether medical mask use by people in a community setting with ILI…" Generalization of data from this study to health care facilities would need significant consideration of variables that differ between settings, which impact transmission (e.g., cues for compliance, proximity of other susceptible individuals [patients, HCW, visitors], ventilation dynamics of the room, adherence to other contact precautions such as hand hygiene). Nevertheless, that discussion would be beyond the scope of this paper.
Authors' reply: We have added following statement "… whether medical mask use by people in a community setting with ILI…". Please see page 6 para 2. We have reduced the discussion around the implications in healthcare settings.

6. Page 6, Line 2: 225 index cases included (also mentions 225 in abstract), but 245 are noted later in the paragraph (123+122) and in Figure 1. Please correct.
Authors' reply: We have corrected the number. Please see abstract and page 6 para 1.

7. Page 6, Line 3-5: Please provide greater detail on methods for sequence generation.
Authors' reply: We have added following statement, "A research team member (YZ) did the random allocation sequence using Microsoft Excel and doctors enrolled the participants randomly to intervention and control arms. Patients had an equal chance to be in the either intervention or control arm". Please see page 7 last para.

8. Page 6, line 14: In consideration of the cultural context in Beijing, is there any risk of contact transmission during meal times wherein index cases may have shared use of utensils or are likely to have contaminated shared food sources (e.g., 'buffet-style' dining?) that is necessary to note?
Authors' reply: We agree it is possible that transmission may have occurred during meal times. This would have the effect of biasing the results toward the null. We have added following statement in limitation section, "It is possible that infection transmission may have occurred during meal times (when patients were not required to wear a mask). This would have the effect of biasing the results toward the null". Please see page 17 last para. We have added a statement in the summary as well.

9. Page 6, line 15: Was proper hand hygiene technique shown?
Authors' reply: We did not observe hand hygiene technique in this study – it was based on self-report.

10. Page 6-7, Outcome measures: Would you clarify who performed daily measurements?
Authors' reply: Index cases were asked to report if any symptoms appeared in household members. Staff in the district CDC also contacted index cases via telephone on every second day to check whether any household member developed symptoms.

We have modified the following para, “Symptoms in the family members were also recorded in the diary cards and index cases were asked to report any symptom. The index cases were asked to contact the study coordinator if any of the following symptoms appeared in household members: cough, nasal congestion, runny nose, sore throat, sneezes, chill, lethargy, loss of appetite, abdominal pain and muscle or joint aches. The study coordinator then assessed the household member and completed a follow-up survey. Samples obtained from all symptomatic cases. All index cases in the intervention and control arms were also asked to document compliance with mask use. Diary cards to record mask use were given to each index case, and they were asked to carry them during the day. Diary cards were returned to the investigators at the end of the study. Staff in the district CDC also contacted index cases via telephone on every alternate day to check whether any household member developed symptoms”. Please see page 9 last para and page 10 first para.

11. Were index cases required to report daily whether or not household contacts had symptoms? Or was no notice from index cases assumed to be indication of no household contact having symptoms? If there are no data being collected daily except when a notification occurs, then is this really “Primary endpoints measured in household contacts on a daily basis”? (It appears that some answers appear in the “Data collection and follow-up” section, but clarity sooner would be better.)
Authors’ reply: We have removed “on a daily basis” from the outcome measure. Please see page 8 last para. Index cases were required to report when household contacts become symptomatic. Staff in district CDC also contacted index cases via telephone on every second day to check whether any household member developed symptoms. Please see point number 10 above for further detail.

12. Page 6-7, outcome measures: CRI is a very broad definition that would be inclusive of ILI. CRI was used as an outcome – one that was significant in post-hoc analysis; yet, CRI was not used as an eligibility criterion for selecting index cases. Would you add rationale for the use of CRI as an outcome, but not for eligibility? To my knowledge, CRI is not a standard outcome other than having been used in previous publications by your colleagues. Because ILI is used for selection, I feel that it would be more balanced to include in the abstract, in addition to CRI post-hoc analysis data that is already mentioned, the post-hoc analysis data from the results section related to ILI comparison of mask group vs no mask group.
Authors’ reply: We desired high specificity for selection of index cases (to ensure maximal efficiency of measuring infection transmission in the household) and high sensitivity for identifying sick household contacts (to ensure maximal identification of incident infections in contacts). We have added following statement in the methods section, “ILI was used as a selection criterion to achieve high specificity for index cases”. Please see page 7 para 1.
We have added the post-hoc analysis data from the results section related to ILI in the abstract, “……a significantly protective effect against clinical respiratory illness (RR 0.24, 95% CI 0.06 to 0.93), but not against influenza like illness (RR 0.16, 95% CI 0.01 – 1.81) and laboratory confirmed viral infection (RR 0.10, 95% CI 0.01 – 3.32)”. Please see page 2.

13. Page 8, Data collection and follow-up: Regarding accuracy of CRI and ILI outcome measurements: When were diary cards collected? Were symptoms of household members assessed by index cases, then confirmed by a (medically) trained project member during swabs for laboratory detection?
Authors’ reply: Diary cards were collected at the end of the study. The index cases were asked to contact the study coordinator if symptoms appeared in the household members. The study coordinator then assessed the household members and completes a follow-up survey. Samples were also obtained from all symptomatic household contacts. We have added a para to explain this. For detail please see point number 10 above.

14. Page 10, Line 10-13: Regarding the validity for mask and no-mask group assignment for the post-hoc analysis: Were compliance measures for mask use or lack of mask use measured in the control

group? How was the mask use identified (by diary or by interview)? How many times must a mask have been used to be eligible for inclusion in the “mask group” of the post-hoc analysis?
Authors’ reply: All index cases (intervention and control arms) recorded mask use through diary cards. We have modified statement to, “All index cases in the intervention and control arms were also asked to document compliance with mask use”. Please see page 10 para 1.
All cases in the control arm who used a mask, were included in the “mask group” in post-hoc analysis. On average, these cases in the control arm used mask for more than one hour per day.

15. Page 11, Line 20-21: Please add the number of index cases that did NOT use a mask during the trial period that were from the mask arm.
Authors’ reply: Seven index cases from the masks arm did not use a mask during the trial period and were included in the “no mask group” in post-hoc analysis. We have added this number. Please see page 12 last para.

REVIEWER 3

Abstract, Strengths & Limitations:
1. Which of these are strengths and which are limitations?
Authors’ reply: We have edited strengths and limitations. Please see page 4.

2. Pt 2: This point is confusing and poorly written. Please revise for clarity.
Authors’ reply: We have deleted the point 2 and updated the strengths and limitations section. Please see page 4.

Introduction

3. Lines 9-11: Medical masks are not worn in healthcare settings to protect healthcare workers in the U.S. That is the function of a respiratory protection device. The purpose of a surgical or medical mask is the protection of patients from healthcare worker emissions.
Authors’ reply: We have modified sentence to, “by well healthcare workers (HCWs) to protect them from splash and spray of blood and body fluids”. Please see page 5 para 1.

4. Paragraph 2: This paragraph is confusing – it mixes studies evaluating the effectiveness of masks worn by patients with effectiveness of masks worn by healthcare workers. I suggest separating into two different paragraphs focusing on these two different goals.
Authors’ reply: We have modified paragraph to, “…….Among the 9 randomized controlled trials (RCTs) in household and community settings to date, only one examined the role of masks as “source control” and was inconclusive. In other clinical trials, masks were either used by both sick patients (index cases as “source control”) and their household members or used only by household members. Most of these studies failed to show any efficacy of mask use in preventing spread of infections from the sick individuals. Please see page 5, para 2.
“Masks are also used to prevent surgical site infections in the operating theatre (OT), although most studies failed to show any efficacy against this indication. Only one clinical trial reported high infection rates after surgery if masks were not used by the surgeon in the OT. Among the five clinical trials in healthcare setting to see the efficacy of masks/ respirators as respiratory protection, none examined the use of masks as source control”. Please see page 5, para 3.

5. Lines 16-17: There have been at least 3 clinical trials of the effectiveness of medical masks on preventing surgical wound infections – one of these is cited (Tunevall). The others are:
o Pippin DJ, Verderame RA, Weber KK. Efficacy of face masks in preventing inhalation of airborne contaminants. J Oral Maxillofac Surg 1987;45:319-23.
o Davies KJ, Herbert AM, Westmoreland D, Bagg J. Seroepidemiological study of respiratory virus

infections among dental surgeons. Br Dent J 1994;176:262-5.
Authors' reply: We have added discussion on use mask in the operating theatre and added following para, "Masks are also used to prevent surgical site infections in the operating theatre (OT), although most studies failed to show any efficacy against this indication. Only one clinical trial reported high infection rates after surgery if masks were not used by the surgeon in the OT. Among the five clinical trials in healthcare setting to see the efficacy of masks/ respirators as respiratory protection, none examined the use of masks as source control". Please see page 5, para 3. References have also been added.

6. Lines 43-48: Masks worn by whom and infection in whom? (see general note on Paragraph 2 above).
Authors' reply: We have modified sentence to, "Laboratory studies generally support the use of medical masks to prevent spread of infections from influenza and TB patients to their contacts". Please see page 6 para 1.

Patient Involvement (Page 7)
7. This section is confusing, particularly the part related to publishing in an open journal.
Authors' reply: We have removed this section. Please see page 9 para 2.

Discussion
This section is generally well-written and organized, particularly where the discussion is focused on household and community trials. I think the two paragraphs on pages 13-14 and the final paragraph in this section that discuss mask use in healthcare settings should be minimized or deleted altogether. I think there is little to be gained from drawing comparisons between community studies of sick people wearing masks and healthcare settings with either patients or workers wearing masks. The settings and situations are too different.
Authors' reply: We have reduced the discussed around source control in healthcare setting and restructured the discussion.

REVIEWER: 4
1. ITT (main analysis) Vs. the post-hoc analyses - In the abstract the results for the post-hoc analyses should be clearly labelled as such so that the reader is aware of potential biases. Similarly, in the main text the main and the post-hoc analyses are intertwined these should be clearly separated.
Authors' reply: We have added following sentence in the methods section of the abstract, "A total of 43 index cases in the control arm also used a mask during the study period, so a post-hoc analysis was carried out to compare outcomes among household members of index cases who used a mask, with those of index cases who did not use a mask". Following sentence has been modified in the result sections of the abstract, "A post-hoc comparison between the mask versus no-mask groups showed a significantly protective effect against clinical respiratory illness (RR 0.22, 95% CI 0.06 to 0.86), but not against influenza like illness (RR 0.18, 95% CI 0.02 – 1.73) and laboratory confirmed viral infection (RR 0.11, 95% CI 0.01 – 4.40)".
2. The post-hoc analyses are potentially biased as they break the randomisation. A CACE analysis could have been a better alternative [Connell AM. Employing Complier Average Causal Effect Analytic Methods to Examine Effects of Randomized Encouragement Trials. The American journal of drug and alcohol abuse. 2009;35(4):253-259. doi:10.1080/00952990903005882.] -It is not clear how time was measured based on the outcome measures section. As it is currently presented the mention of Cox regression and Kaplan Meier curves in the "Data analysis" section is questionable. Also, I wonder whether the one can afford a multivariable regression model given the very few events encountered. As a rule of thumb, one needs at least ten events per explanatory variable for a reliable estimate. Even when this is relaxed, at least five events were needed. [Vittinghoff E and McCulloch CE. Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression Am. J. Epidemiol. (2007) 165 (6): 710-718 first published online December 20, 2006 doi:10.1093/aje/kwk052]. This

could explain the reported wide confidence intervals.
Authors' reply: We agree with the reviewer that the post-hoc analyses are potentially biased due to loss of randomisation. The main aim of post-hoc analyses was to check the sensitivity of intention-to-treat (ITT) analysis. We have added the following sentence in the limitations section, "Post-hoc analyses are potentially biased due to loss of randomisation and it was added as a sensitivity analysis in this study because of deviations from protocol in mask wearing". Please see page 18 para 3.
Time was measured though the use of diary cards. We have added the following details, "Symptoms in the family members were also recorded in the diary cards and index cases were asked to report if any symptoms appears", and "Data of diary cards were used to calculate person-days of infection incidence". Page 12 para 1.
We agree with the reviewer's comment that a multivariable Cox model was not appropriate due to small number of events. We updated the analyses and reported results from univariable Cox model for ILI and laboratory confirmed influenza. Age of household contact was significantly associated with CRI in a univariable Cox model. Thus only this variable was included in a multivariable Cox model with CRI as there are ten events with this outcome.
We have updated the method and results sections accordingly and added following statements, "Due to very few outcome events encountered a multivariable Cox model was not appropriate. We checked the effect of individual potential confounders on the outcome variable fitting univariable Cox models. Because there are 10 cases of CRI, we included this variable in a multivariable cluster adjusted Cox model. Multivariate analyses were not performed for ILI and laboratory confirmed influenza because of low numbers (Page 12 para 1). ……."In a univariable Cox model only the age of household contact was significantly associated with the CRI (Table 3)". Please see page 14 para 2.
Due to revised analysis the RR and CI values have slightly changed but there is not major difference.
3. Page 10 Lines 13/15 this sentence needs revisiting as it is not clear why you would use information from a multiple regression type model to inform univariate analysis.
Authors' reply: We have modified statement to, "Due to very few outcome events encountered a multivariable Cox model was not appropriate. Please see page 12 para 1.
4. The use of person-day rates should be clearly stated in the analysis section. Also, were rates used in the sample size calculation or were these based on proportions? On a related note, the authors mention on Page 6 Line 43/45 that participants were informed that they could stop wearing a mask once their symptoms have resolved, how was this documented and how did they account for this when calculating person-day data.
Authors' reply: Person-day rates were calculated from the diary card data. Data on masks use was also present in the diary cards and we accounted for this when calculating person-day data. We have added following detail, "Symptoms in the family members were also recorded in the diary cards and index cases were asked to report if any symptoms appeared" (page 19 last para), and "Data from diary cards were used to calculate person-days" (Page 12 para 1).
5. A Cluster RCT: It should be clear in the title that this is a cluster randomised trial.
Authors' reply: We have added this.
6. It is not clear whether the main analyses were made at the individual level or at the cluster (household) level. It should be clearly stated in this case how clustering was adjusted for in a similar fashion to how this was mentioned for the Cox regression.
Authors' reply: We have added following statement in the analysis section, "The analyses were conducted in individual level and hazard ratios (HR) were calculated using Cox proportional hazards model after adjusting for clustering by household by adding a shared frailty to the model." Please see page 12 para 1.
7. The modified CONSORT for cluster randomised trial should have been used.
Authors' reply: We have attached modified CONSORT list for cluster randomised trial.
8. Reporting of ICCs is advised to aid sample size calculations for future studies.
Authors' reply: We have added ICCs. Please see table 2.
9. Page 11 Line 11/15 a clarification whether these were 4 members from 4 different households or whether some were from the same household should be clarified. Ditto for the ensuing sentence

Authors' reply: 4 members were from 4 different households.
Minor Points:
10. In the discussion, reasons for the study being under-powered are worthy of discussion in order to inform larger trials. Also it would be interesting to discuss the discrepancy between the observed proportions/rates of the main outcomes in this study and the assumed ones used in the sample size calculations.
Authors' reply: We have added following statement in the limitation section, "In the sample size calculations, we assumed 16-20% attack rate of CRI in the control arm, based on the results of a previously published household mask trial. However the secondary attack rates were much lower in this study which might be due to testing only symptomatic cases." Please see page 17 last para.
11. Table 1 should only include baseline data. Post randomisation data should be presented in a separate table, such as "average hour mask wearing".
Authors' reply: We have removed post randomisation data from the table 1.
12. P13 L15 replace "to examine to" with "to examine the"
Authors' reply: We have edited this.

VERSION 2 – REVIEW

REVIEWER	Gary garber and Jeffrey Smith Public Health Ontario Canada
REVIEW RETURNED	06-Sep-2016

GENERAL COMMENTS	The authors have satisfactorily responded to our previous comments. Some additional comments are as follows: General The article needs to be combed through for other grammar mistakes and consistency issues of which there are many. E.g., in the strengths and limitations second the first bullet has an extra period at the end; the references contain a mixture of sentence and title case structure as well as in ref 10 there is a lack of punctuation where needed; some headings are title case and other sentence case. Hyphenation throughout article should be added: influenza-like illness; laboratory-confirmed viral respiratory infection. Choose one or the other: "follow up" vs "follow-up". The authors use laboratory-confirmed viral infections throughout. Although more verbose it is more accurate to state "laboratory-confirmed viral respiratory infections". Source control is sometimes placed in quotations. I believe this is a standard term that does not require quotations in any instance of the article. Interchange of household members and family members is confusing. Household members could mean to include roommates or borders. Family members is more restrictive. Contacts are also used to describe the household members of index cases. Specific 1. Page 2, intervention section of abstract: The second sentence is confusing. Probably requires a period between "study period" and "a

	post-hoc". For clarity the second sentence could be written as, "An as-treated post-hoc analysis was performed by comparing outcomes among household members of index cases who used a mask (mask group) with household members of index cases who did not use a mask (no-mask group)." 2. Page 2, main outcome measure section of abstract: Suggest rewrite as, "Primary outcomes measured in household members were clinical respiratory illness, ILI, and laboratory-confirmed viral respiratory infection." Reason to rewrite: Remove "or bacterial". Authors now mention laboratory-confirmed bacterial respiratory infection as an outcome (previously omitted). The explanation for the deviation from protocol has been added to the CONSORT checklist located as a supplemental file. Therefore, does not need be included in the article. Also, add a period to the end of this paragraph. 3. Page 2, results section of abstract: The ITT values reported are the hazard ratios (HR) and not the relative risks (RR). Please either use the RR values or change the abbreviations to HR. 4. Page 4, point 1: add "cluster" for the design type, and add "medical" for the mask type to differentiate from other types of masks (c loth/paper). 5. Page 4, point 4: Abstract definitively states study "was" underpowered, but here it is not absolute: "may have been". The study was underpowered (lower attack rates than anticipated for sample size calculation). The same "may have been" is also in the discussion section page 16. 6. Page 5, paragraph 1: I disagree with the removal of medical masks (similarly termed surgical masks, [medical] procedure masks, isolation masks, laser masks, fluid-resistant masks and face masks in the U.S.) as not being a common use of medical masks in healthcare settings as noted by Reviewer 3. FDA approved medical masks are regulated by the same standards as surgical masks (http://www.fda.gov/RegulatoryInformation/Guidances/ucm072549.htm) and are frequently used as a form of respiratory protection rather than respirators for droplet transmission-based respiratory infections. I would leave this to the editor's discretion. 7. Page 5, paragraph 1: The wording that cloth and medical masks are developed to "prevent the spread of infection from the wearer in operating theatres" is not quite accurate because one would hope that the wearer performing an operation does not have an infection. Understandably, opportunistic infections are possible from normal oral commensals from a healthy (well) surgeon in an operating theatre. Therefore infection is not spread. Instead, this should be more accurately described as surgical site infections (used in second clause of the sentence), or "…source control to prevent contamination of sterile sites by the wearer in operating theatres (OTs)…". 8. Page 6, paragraph on Design: Remove the sentence about the duration of recruitment. The recruitment period given after that line is more than 6 weeks. 18 Nov – Jan 20 ~9 weeks. 9. Page 7, last line of first paragraph: "with less than two other people and have other ill household members" should use "or" to join the list of items and a serial comma should be used after "other people". 10. Page 8, top of page: The authors could refer to removal of the mask as "doffing the mask" if the donning term is used. 11. Page 8, last sentence of top paragraph: Was mask use by other household members recorded? If not, please modify to "… was not required and not reported." 12. Page 8, 2nd paragraph: Please write "Respiratory illness

	outcomes…" as the lead for the paragraph to be consistent with the intervention section paragraph that says "respiratory illness was measured in household contacts". 13. Page 8, last paragraph: In author's response it says phone calls were made every second day, but in the description it is twice-weekly. Page 9 has a description that CDC staff called every other day. How many people were contacting these participants to collect data, which data was used, or how was it handled from the two sources? Is it relevant to differentiate where your symptom data was obtained from? (CDC staff versus study coordinator.) 14. Page 9, first paragraph: The first sentence is worded awkwardly and redundantly. Suggest: "At baseline detailed clinical and demographic information including household structure was collected from index cases and their household members." 15. Page 9, second paragraph: "Samples obtained from all…" add "were" to sentence. 16. Page 10, last paragraph: Abbreviation (NAT) was already defined on page 8. Do not need to define again here. 17. Page 10, last paragraph: Remove the last section about PCR for bacterial species because authors stated that this was not conducted due to financial constraints. If it was conducted, please report the data and update the CONSORT statement. 18. Page 11, data analysis paragraph: "Relative risks were calculated for the mask group." Do you mean the mask arm? Mask group is the post-hoc analysis. 19. Page 12, top paragraph: ".. and laboratory confirmed influenza.." Do you mean laboratory confirmed viral respiratory infection? This error is also made in the last paragraph on page 16. 20. Page 12, second paragraph: Please define the specific criteria of mask use for the 43 index cases in the control arm that wore a mask that made them eligible for the mask group post-hoc analysis (e.g., at least once a day for 1 hour), and the specific criteria for lack of mask use of the 7 index cases in the mask arm that made them eligible for the no-mask group post-hoc analysis (e.g., no more than once a day and less than 1 hour per day). Though not a necessity, please also add any rationale to support the criteria used. 21. Page 13, paragraph 2: "Two laboratory confirmed… from respective index case." Suggested to rewrite as: "Two laboratory-confirmed viral respiratory infections were identified among symptomatic household members from separate households. One household member had the same infection (influenza H1N1) as the respective index case. Rhinovirus was isolated from the other household member. However, no pathogen was isolated from respective index case." 22. Page 13, paragraph 2: Rewrite second last sentence to "The two cases of laboratory confirmed viral respiratory infections of household members occurred in separate study arms (RR 0.97, 95% CI 0.06 to 15.5)." 23. Page 14, discussion paragraph: "rates of secondary infections in household members were consistently lower" is not true. This should state rates of CRI and ILI. Laboratory confirmed viral respiratory infections was essentially identical (1/123 vs 1/122). 24. Page 14, discussion paragraph: "additional analysis by actual mask used showed significantly lower rates of clinical infection…" Again this is not true. Please be careful with wording. The rates of clinical respiratory illness were significantly lower. 25. Page 15, discussion about Canini study: The OR is reported as 0.95, but the intervention had a higher incidence of ILI than control arms (16.2% vs 15.8%). Normally a lower OR favours intervention. Would you please check that the values you have used are correct

	and maybe recalculate the OR to make more sense with respect to standard conventions? 26. Would you please comment on the removal of the hand washing, average hour of home stay, average hour of contact, and average hour of mask wearing from Table 1? It would seem this is important information to consider for the applicability of the data presented: e.g., 16.6 hours of home stay, 10.4 hours of contact, but only 4.4 hours of mask use in the intervention arm seems to indicate a long period of potential exposures for household members to the index case.

REVIEWER	Lisa M Brosseau University of Illinois at Chicago United States I have not conducted research with these investigators, but have participated in grant submissions that have not been funded to date.
REVIEW RETURNED	05-Sep-2016

GENERAL COMMENTS	I believe the authors have been responsive to my and other reviewers' comments.

REVIEWER	Mona Kanaan Senior Lecturer University of York, UK
REVIEW RETURNED	07-Sep-2016

GENERAL COMMENTS	I would like to thank the authors for addressing my earlier comments. This is a much improved version. Some points, mainly minor, remain to be addressed. • In the abstract under results and discussion, the post-hoc results should be de-emphasised. Readers should also be alerted to the fact that there were very few events that these statistics are based on. • Could the authors provide the mean time for wearing a mask by arm? • There are lots of variables that were collected in the diaries but not reported here. Are the authors going to provide supplemental files for these analyses or are they going to be reported in another publication? • Please add a reference for STATA 13 and cite. • Under the RESULTS section Page 12 Line 46, specify what the “Some differences” are. • Under Table 1 there is a note next to the first dagger symbol (†), it is not clear to what it refers to. Also, the first variable under Household should be mean number of household members. • The quality of the Kaplan Meier curves could be improved. It should be clarified that the scale used represents only a fraction of the 0-1 range.

	• In Tables 3 and 5, it should be clarified why only age is being reported in these tables as it is the case in the main text.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

General

1. The article needs to be combed through for other grammar mistakes and consistency issues of which there are many. E.g., in the strengths and limitations second the first bullet has an extra period at the end; the references contain a mixture of sentence and title case structure as well as in ref 10 there is a lack of punctuation where needed; some headings are title case and other sentence case.
Authors' reply: We have checked the paper and have made suggested changes in text and references.

2. Hyphenation throughout article should be added: influenza-like illness; laboratory-confirmed viral respiratory infection. Choose one or the other: "follow up" vs "follow-up".
Authors' reply: We have made suggested changes and used following words throughout the paper; influenza-like illness, laboratory-confirmed viral respiratory infection and "follow-up.

3. The authors use laboratory-confirmed viral infections throughout. Although more verbose it is more accurate to state "laboratory-confirmed viral respiratory infections".
Authors' reply: We have made suggested changes in the paper.

4. Source control is sometimes placed in quotations. I believe this is a standard term that does not require quotations in any instance of the article.
Authors' reply: We have made suggested changes and removed quotations.

5. Interchange of household members and family members is confusing. Household members could mean to include roommates or borders. Family members is more restrictive. Contacts are also used to describe the household members of index cases.
Authors' reply: The term "household" had been commonly used in previous studies. We have used the same term at most places. We have made suggested changes and changed the term "family member" to "household member" throughout the paper.

Specific

1. Page 2, intervention section of abstract: The second sentence is confusing. Probably requires a period between "study period" and "a post-hoc". For clarity the second sentence could be written as, "An as-treated post-hoc analysis was performed by comparing outcomes among household members of index cases who used a mask (mask group) with household members of index cases who did not use a mask (no-mask group)."
Authors' reply: We have added following sentence, "An as-treated post-hoc analysis was performed by comparing outcomes among household members of index cases who used a mask (mask group) with household members of index cases who did not use a mask (no-mask group)". Please see the abstract.

2. Page 2, main outcome measure section of abstract: Suggest rewrite as, "Primary outcomes measured in household members were clinical respiratory illness, ILI, and laboratory-confirmed viral respiratory infection." Reason to rewrite: Remove "or bacterial". Authors now mention laboratory-

confirmed bacterial respiratory infection as an outcome (previously omitted). The explanation for the deviation from protocol has been added to the CONSORT checklist located as a supplemental file. Therefore, does not need be included in the article. Also, add a period to the end of this paragraph.
Authors' reply: We have modified sentence to, "Primary outcomes measured in household members were clinical respiratory illness, ILI, and laboratory-confirmed viral respiratory infection". Please see the abstract.

3. Page 2, results section of abstract: The ITT values reported are the hazard ratios (HR) and not the relative risks (RR). Please either use the RR values or change the abbreviations to HR.
Authors' reply: ITT values in table 2 are reported as relative risk (RR). Survival analyses in table 3 and 5 are reported as hazard ratios (HR).

4. Page 4, point 1: add "cluster" for the design type, and add "medical" for the mask type to differentiate from other types of masks (cloth/paper).
Authors' reply: We have modified sentence to, "We conducted a cluster randomised control trial to examine the efficacy of medical masks as "source control". Please see page 4.

5. Page 4, point 4: Abstract definitively states study "was" underpowered, but here it is not absolute: "may have been". The study was underpowered (lower attack rates than anticipated for sample size calculation). The same "may have been" is also in the discussion section page 16.
Authors' reply: We have modified sentence to, "The study indicates a potential benefit of medical masks for source control, but was limited by small sample size and low secondary attack rates. Larger trials are needed to confirm efficacy of medical masks as source control". Please see the abstract. Same change has made in the discussion as well. Please see page 15, first para.

6. Page 5, paragraph 1: I disagree with the removal of medical masks (similarly termed surgical masks, [medical] procedure masks, isolation masks, laser masks, fluid-resistant masks and face masks in the U.S.) as not being a common use of medical masks in healthcare settings as noted by Reviewer 3. FDA approved medical masks are regulated by the same standards as surgical masks (http://www.fda.gov/RegulatoryInformation/Guidances/ucm072549.htm) and are frequently used as a form of respiratory protection rather than respirators for droplet transmission-based respiratory infections. I would leave this to the editor's discretion.
Authors' reply: We have modified sentence to, "Medical masks are commonly used in healthcare settings for two main purposes: 1) by well healthcare workers (HCWs) to protect them from infections transmit by droplet route and splash and spray of blood and body fluids; and ……..". Please see page 5 para 1.

7. Page 5, paragraph 1: The wording that cloth and medical masks are developed to "prevent the spread of infection from the wearer in operating theatres" is not quite accurate because one would hope that the wearer performing an operation does not have an infection. Understandably, opportunistic infections are possible from normal oral commensals from a healthy (well) surgeon in an operating theatre. Therefore infection is not spread. Instead, this should be more accurately described as surgical site infections (used in second clause of the sentence), or "…source control to prevent contamination of sterile sites by the wearer in operating theatres (OTs)…".
Authors' reply: We have modified the sentence to, "Cloth and medical masks were originally developed as source control to prevent contamination of sterile sites by the wearer in operating theatres (OTs)". Please see page 5 para 1.

8. Page 6, paragraph on Design: Remove the sentence about the duration of recruitment. The recruitment period given after that line is more than 6 weeks. 18 Nov – Jan 20 ~9 weeks.
Authors' reply: We have removed this sentence. Please see page 6 para 3.

9. Page 7, last line of first paragraph: “with less than two other people and have other ill household members” should use “or” to join the list of items and a serial comma should be used after “other people”.
Authors’ reply: We have modified the sentence to, “Patients who were unable or refused to give consent, had onset of symptoms >24 hours prior to recruitment, were admitted to hospital, resided in a household with less than two other people, or have other ill household members at home were excluded from the study. Please see page 7 para 1.

10. Page 8, top of page: The authors could refer to removal of the mask as “doffing the mask” if the donning term is used.
Authors’ reply: We have modified the sentence to, “Index cases were shown how to wear the mask and instructed to wash their hands when donning and doffing the mask”. Please see page 8 para 1.

11. Page 8, last sentence of top paragraph: Was mask use by other household members recorded? If not, please modify to “… was not required and not reported.”
Authors’ reply: We have modified the sentence to, “Mask use by other household members was not required and not reported”. Please see page 8 para 1.

12. Page 8, 2nd paragraph: Please write “Respiratory illness outcomes…” as the lead for the paragraph to be consistent with the intervention section paragraph that says “respiratory illness was measured in household contacts”.
Authors’ reply: We have modified the sentence to, “Respiratory illness outcomes were measured in household contacts of the index cases”. Please see page 8 para 2.

13. Page 8, last paragraph: In author’s response it says phone calls were made every second day, but in the description it is twice-weekly. Page 9 has a description that CDC staff called every other day. How many people were contacting these participants to collect data, which data was used, or how was it handled from the two sources? Is it relevant to differentiate where your symptom data was obtained from? (CDC staff versus study coordinator.)
Authors’ reply: CDC staff and study coordinator is the same person. We have modified the statement to. “The study coordinator also contacted index cases via telephone on every alternate day to check whether any household member developed symptoms”. Please see page 9 last para.
We have deleted following statement, “The study coordinator also performed twice-weekly follow-up phone calls to the families to actively ascertain incident illness in household members”. Please see page 8 last para.

14. Page 9, first paragraph: The first sentence is worded awkwardly and redundantly. Suggest: “At baseline detailed clinical and demographic information including household structure was collected from index cases and their household members.”
Authors’ reply: We have modified the sentence to, “At baseline detailed clinical and demographic information including household structure was collected from index cases and their household members”. Please see page 9 para 1.

15. Page 9, second paragraph: “Samples obtained from all…” add “were” to sentence.
Authors’ reply: We have modified the sentence to, “Samples were obtained from all symptomatic cases”. Please see page 9 para 2.

16. Page 10, last paragraph: Abbreviation (NAT) was already defined on page 8. Do not need to define again here.
Authors’ reply: We have edited this. Please see page 10, last para.

17. Page 10, last paragraph: Remove the last section about PCR for bacterial species because

authors stated that this was not conducted due to financial constraints. If it was conducted, please report the data and update the CONSORT statement.
Authors' reply: We have deleted following sentence, "NAT using a multiplex PCR was also done on the same DNA/RNA extract as used for the viral PCR (Seegen, Inc., Seoul, Korea) for Streptococcus pneumoniae, Mycoplasma pneumoniae, Bordetella pertussis, legionella, chlamydia and Haemophilus influenzae type B". Please see page 10, last para.

18. Page 11, data analysis paragraph: "Relative risks were calculated for the mask group." Do you mean the mask arm? Mask group is the post-hoc analysis.
Authors' reply: We have modified the statement to, "Relative risks were calculated for the mask arm". Please see page 11, last para.

19. Page 12, top paragraph: ".. and laboratory confirmed influenza.." Do you mean laboratory confirmed viral respiratory infection? This error is also made in the last paragraph on page 16.
Authors' reply: We have changed it to "laboratory-confirmed viruses". Please see page 12 para 1 and page 17 para 1.

20. Page 12, second paragraph: Please define the specific criteria of mask use for the 43 index cases in the control arm that wore a mask that made them eligible for the mask group post-hoc analysis (e.g., at least once a day for 1 hour), and the specific criteria for lack of mask use of the 7 index cases in the mask arm that made them eligible for the no-mask group post-hoc analysis (e.g., no more than once a day and less than 1 hour per day). Though not a necessity, please also add any rationale to support the criteria used.
Authors' reply: We have modified the sentence to, "A total of 43 index cases in the control arm also used a mask during the study period (at least one hour per day) and 7 index cases in the masks arm did not use a mask at all,……..". Please see page 12, para 2.

21. Page 13, paragraph 2: "Two laboratory confirmed… from respective index case." Suggested to rewrite as: "Two laboratory-confirmed viral respiratory infections were identified among symptomatic household members from separate households. One household member had the same infection (influenza H1N1) as the respective index case. Rhinovirus was isolated from the other household member. However, no pathogen was isolated from respective index case."
Authors' reply: We have modified the statement as suggested by the reviewer. Please see page 13 last para.

22. Page 13, paragraph 2: Rewrite second last sentence to "The two cases of laboratory confirmed viral respiratory infections of household members occurred in separate study arms (RR 0.97, 95% CI 0.06 to 15.5)."
Authors' reply: We have modified the statement as suggested by the reviewer. Please see page 13 last para.

23. Page 14, discussion paragraph: "rates of secondary infections in household members were consistently lower" is not true. This should state rates of CRI and ILI. Laboratory confirmed viral respiratory infections was essentially identical (1/123 vs 1/122).
Authors' reply: We have modified the statement to, "We did not find a significant benefit of medical masks as source control, but rates of CRI and ILI in household members were consistently lower in the mask arm compared to the control arm". Please see page 15 para 1.

24. Page 14, discussion paragraph: "additional analysis by actual mask used showed significantly lower rates of clinical infection…" Again this is not true. Please be careful with wording. The rates of clinical respiratory illness were significantly lower.
Authors' reply: We have modified the statement to, "The additional analysis by actual mask use

showed significantly lower rates of CRI in mask group compared to the no-mask group, ……”. Please see page 15 para 1.

25. Page 15, discussion about Canini study: The OR is reported as 0.95, but the intervention had a higher incidence of ILI than control arms (16.2% vs 15.8%). Normally a lower OR favours intervention. Would you please check that the values you have used are correct and maybe recalculate the OR to make more sense with respect to standard conventions?
Authors' reply: We have deleted OR and have added mean difference in the attack rates between intervention and control arms, i.e. mean difference 0.40%, 95%CI: -10% to 11%, P= 1.00. The Odds for the interventional arm may be lower owing to adjustment for other factors which were significantly associated with the ILI, such as mild vs sever symptoms, elevated temperature, age and gender of household contact.

26. Would you please comment on the removal of the hand washing, average hour of home stay, average hour of contact, and average hour of mask wearing from Table 1? It would seem this is important information to consider for the applicability of the data presented: e.g., 16.6 hours of home stay, 10.4 hours of contact, but only 4.4 hours of mask use in the intervention arm seems to indicate a long period of potential exposures for household members to the index case.
Authors' reply: These variables were included in the submitted version but removed after suggestion of the other reviewer during first review. We have again added data in the text in result section. Perhaps the editor can make a final decision on what variables to include. We have added that, "Duration of contact of index cases with household members was 10.4 hours and 11.1 hours in mask and control arms respectively. On average, participants in the mask arm used a mask for 4.4 hours, while participants in the control arm used a mask for 1.4 hours". Please see page 13 last para.
We agree, and this reflects low compliance as seen in other mask studies. We have mention in the intervention section that, "Index cases were asked to wear a mask (3M 1817 surgical mask) whenever they were in the same room as a family member or a visitor to the household". We have provided the data as reported by participants, as per intention-to-treat analysis. This is one of the reasons we added the post-hoc analysis.

Reviewer: 4
Reviewer Name: Mona Kanaan
Institution and Country: Senior Lecturer, University of York, UK Competing Interests: NONE

1. In the abstract under results and discussion, the post-hoc results should be de-emphasised. Readers should also be alerted to the fact that there were very few events that these statistics are based on.
Authors' reply: We have modified sentence in results section to," A post-hoc comparison between the mask versus no-mask groups showed a protective effect against clinical respiratory illness but not against ILI and laboratory-confirmed viral respiratory infections".
We have modified conclusion to, "The study indicates a potential benefit of medical masks for source control, but was limited by small sample size and low secondary attack rates. Larger trials are needed to confirm efficacy of medical masks as source control." Please see the abstract.
2. Could the authors provide the mean time for wearing a mask by arm?
Authors' reply: We have added mean time for wearing a mask by arm in the text. We have added following statement, "On average, participants in the mask arm used a mask for 4.4 hours, while participants in the control arm used a mask for 1.4 hours". Please see page 13, last para.

3. There are lots of variables that were collected in the diaries but not reported here. Are the authors going to provide supplemental files for these analyses or are they going to be reported in another publication?
Authors' reply: These data are being analysed for a separate paper.

4. Please add a reference for STATA 13 and cite.
Authors' reply: We have added a reference. Please see reference # 30.

5. Under the RESULTS section Page 12 Line 46, specify what the "Some differences" are.
Authors' reply: We have modified statement to, "There was no significant difference between arms, and most characteristics, including medication use (data not shown), were generally similar".

6. Under Table 1 there is a note next to the first dagger symbol (†), it is not clear to what it refers to. Also, the first variable under Household should be mean number of household members.
Authors' reply: This sign was used to indicate variables that were created by taking average hours over the trial period. These variables were removed after last review on request of the other reviewer, but have now again added.

7. The quality of the Kaplan Meier curves could be improved. It should be clarified that the scale used represents only a fraction of the 0-1 range.
Authors' reply: We have provided separate TIF files of all Kaplan Meier curves as well. In the foot note we have mentioned that the scale used in Kaplan Meier curves represents only a fraction of the 0-1 range.

8. In Tables 3 and 5, it should be clarified why only age is being reported in these tables as it is the case in the main text.
Authors' reply: We have added following statements as footnote in tables 3 and 5, "Multivariate analysis was performed as there were 10 cases of CRI and age was also significant in the univariate analysis. Multivariate analyses were not performed for ILI and laboratory-confirmed viral respiratory infections due to low number of cases.

VERSION 3 – REVIEW

REVIEWER	Mona Kanaan University of York
REVIEW RETURNED	28-Oct-2016

GENERAL COMMENTS	Please note that Table 1 and the Figures do not seem to have been updated in the version R2. These need to be updated as per the authors' response.

VERSION 3 – AUTHOR RESPONSE

Reviewer Name: Mona Kanaan
Institution and Country: Senior Lecturer, University of York, UK Competing Interests: NONE
1. Please note that Table 1 and the Figures do not seem to have been updated in the version R2. These need to be updated as per the authors' response.
Authors' reply: We have updated tables and figures. We have added figure legend as well. Please see page 20 and 23.