PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Evaluating spillover of HIV knowledge from study participants to their network members in a stepped-wedge behavioral intervention in Tanzania
AUTHORS	Lienert, Jeffrey; Fawzi, Mary C; McAdam, Keith; Kaaya, Sylvia; Liu, Yuanyuan; Todd, Jim; Andrew, Irene; Onnela, Jukka Pekka

VERSION 1 – REVIEW

REVIEWER	Hank Green Indiana University School of Public Health
REVIEW RETURNED	26-Sep-2019

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. I am keenly interested in the development of HIV interventions that explore peer processes as mechanisms for information transfer and behavior change. While I believe that there are some interesting and important findings presented in this manuscript, I believe that it requires a great deal of revision to address serious concerns. Given the level of revision that should be undertaken, I believe it best that I provide feedback at a very high level at this time. My concerns focus on issues of completeness, clarity, and significance. This paper explores loss to followup and increase in knowledge among network members in a behavioral intervention in Tanzania. The analyses first compare change agents to their network members, then assess variables associated with network member loss to follow up with GEEs, then conduct multivariate regressions to assess loss to followup, then model direct and indirect knowledge transfer from change agents to network members. I believe that this paper should provide a great deal more detail about the design of the study—such as how many network members change agents were asked to enroll-- what the demographics of the change agents and their network members were, should justify the demographic and contextual variables included in the model—such as having a source of water. Etc. all variables used should be justified in the literature review and should be defined in a separate section. The presentation of the study should be far more formal, complete, and systematic than what is provided here. Over and above the exploration of factors associated with loss to followup among network members, we should know what factors were associated with loss to followup for change agents, and whether network members are lost because their related change agents are lost… if this is presented in other publications then it should be summarized here so that we get an introduction that motivates the reasons for the analyses, justifies the variables expected to be important, presents hypotheses regarding those

variables, describes the analytic strategies and justifies any complicated or non-standard analytic approaches, presents results starting with descriptive statistics, then bivariate statistics, then multivariate models, summarizes the key results and then presents discussion. For me, this would also describe the study, describe the change agents, describe the network members including how many on average were recruited per change agent. Assess whether there was bias in recruiting related to characteristics of the change agents, present information on the matched pairs that were included in the analysis, assess whether any of the clusters of dyads were influential enough to require adjustments to the standard errors related to clustering, etc.

Once the study is more fully described, the variables and analyses justified, and the results presented more completely, only then can we begin to assess the significance and the importance of this study and the implications. For now I will only say that I am not convinced that the intervention will only have one-step effects because if participants are motivated to discuss their HIV knowledge with their network members this may also motivate network members to do the same in their own networks. but once all of the details are provided more fully, I may be convinced otherwise.

As it stands now, the manuscript just raises all of these questions because the information is not provided. The manuscript feels incomplete and lacking in key details for readers to fully understand the intervention, the present study, the results and the implications.

REVIEWER	Ashley Buchanan University of Rhod Island, United States
REVIEW RETURNED	14-Oct-2019

GENERAL COMMENTS	General Comments

Overall, this is an interesting article and an important application of methods to evaluate spillover effects of an HIV prevention intervention. The authors evaluate the role of change agents (CAs) in an HIV behavioral educational intervention among people living with or at risk for HIV in Tanzania. Their paper addresses an important gap in the literature, specifically how effective are change agents in spreading informational knowledge to their network members (NMs) in this population. My primary concerns are limited important details about the stepped wedge design, including appropriate adjustment for confounders, and the identification of dyads assumed to be non-overlapping in the analysis. Please find a summary of major comments and detailed minor comments below.

Specific Comments for Revision

Major Comments

The definition of spillover seems narrow, could be clarified and also needs a reference. In the literature, there are other terms for this, including indirect and disseminated effects (Hudgens and Halloran, 2008; Buchanan, 2018). When evaluating the effects of an intervention, spillover is when one participant's exposure to the intervention affects another participant's outcome. I would also suggest acknowledgement of the existing literature about the spillover effects of HIV knowledge due to peer interventions (Latkin, et al. 2009; Ghosh, et al. 2017). The authors claim that spillover

effects of HIV knowledge interventions are unknown, when in fact, there is some evidence among people who inject drugs in the United States. I would suggest this paper provides evidence for spillover effects of HIV knowledge interventions to prevent HIV transmission in Sub-Saharan Africa.

More details are needed to describe the stepped wedge design of the trial and the implications for the analysis of spillover effects. Were dyads analyzed according to their randomized crossover time? How was the stepped wedge design accounted for in the analysis? In the methods, it is not clear if there is an intervention group and a control group, or a single group where all participants were randomized to receive the intervention during a different wave. Based on the Discussion section, it seems that the CAs were randomly assigned to the intervention or no intervention. Please clarify this point.

More details are needed about the change agents. Please define "change agents" the first time this is discussed. How many network members could each CA recruit? Were these connections social only, or could the connections also be HIV risk connections? As currently written, it is not clear if the clusters are dyads or could be larger in the study. By design, was it only one CA and one NM per cluster? Or were dyads created for the analysis based on more complex egocentric networks? If so, how big were the egocentric networks? The process of how dyads were

selected is described in the Discussion Section. Please describe how dyads were ascertained from study data in the Methods Section.

Did you collect any information to look at possible contamination from intervention to control clusters?

The claim that the overlapping egocentric neighborhoods did not impact the analysis assuming the neighborhoods were non-overlapping is a strong assertion. Perhaps provide some insights into why randomly removing dyads is a valid way to test this assumption? If this assumption, does not hold, how would that impact your results?

Why did the authors restrict to a dyad analysis, when methods for egocentric networks exist? Please see Buchanan, et al. (2018) for nonoverlapping egocentric networks and Forestiere, et al. (2019) for possibly overlapping egocentric networks.

For the instrument used to collect knowledge, what was the Chronbach's alpha or Omega in your study sample? When describing the other studies that used this instrument, please be sure to describe the study population for South Africa and in the original study. Is it valid to combine across questions of this instrument to obtain a single measure (complete knowledge vs. incomplete)? Please provide some justification for this.

Is using an intercept only GEE an appropriate test of homophily? Instead, the authors could consider an exponential random graph model to assess homophily and include nodal homophily terms. I did not see the homophily results described the Results Section. Please be sure to include these in the Results Section.

For the logistic regression in the second aim, how did the authors account for correlation in the study? Please clarify what you mean by interval censoring of the network members. Do you mean you defined their drop out date as halfway between their last visit and their subsequent missed visit? Please see the paper by Lesko, et al. about when to censor.

For the third aim, please define what you mean by "the putative causal mechanism". I suggest instead of "spillover analysis", the authors could use "evaluation of spillover effects". In this analysis, the authors seem to conflate mediation with spillover. In the Vanderweele paper cited, there is a true mediator variable in the study (i.e., classroom quality) that is a separate concept from spillover between study participants. Spillover is the effect on the CA's exposure on the other dyad member's HIV knowledge. This is a distinct concept from an exposure affecting an outcome through a mediator. In this paper, there does not seem to be a mediator, but rather only interest in estimating spillover effects. I would suggest the authors estimate the direct, indirect, composite and overall effects as described in Hudgens and Halloran (2008) for clustered data or more specifically for dyads in Vanderweele's textbook (pages 399 – 402).

This analysis uses data from a study that employed a stepped wedge randomized design. If randomization is lacking either at the dyad (or egocentric network) or individual level, there

may be confounding bias when trying to identify the direct, indirect, composite and overall effects using the observed study data. Furthermore, due to the stepped wedge design, there may also be confounding by calendar time. Did the authors adjust for possible confounders in their analysis of spillover?

Would it be possible to adjust for selection bias due to loss to follow-up in your estimators of spillover? If not, this limitation is well covered in the discussion. Is the partial interference assumption made in your analysis? If so, are there any doubts about this assumption? If so, please be sure to include in the Discussion Section.

References:

Buchanan, A. L., Vermund, S. H., Friedman, S. R., & Spiegelman, D. (2018). Assessing individual and disseminated effects in network-randomized studies. American Journal of Epidemiology, 187(11), 2449-2459.

Forastiere, L., Airoldi, E. M., & Mealli, F. (2016). Identification and estimation of treatment and interference effects in observational studies on networks. arXiv preprint arXiv:1609.06245.

Ghosh, D., Krishnan, A., Gibson, B., Brown, S. E., Latkin, C. A., & Altice, F. L. (2017). Social network strategies to address HIV prevention and treatment continuum of care among at-risk and HIV-infected substance users: a systematic scoping review. AIDS and Behavior, 21(4), 1183-1207.

Hudgens, M. G., & Halloran, M. E. (2008). Toward causal inference with interference. Journal of the American Statistical Association, 103(482), 832-842.

Latkin, C. A., Donnell, D., Metzger, D., Sherman, S., Aramrattna, A., Davis-Vogel, A., ... & Celentano, D. D. (2009). The efficacy of a network intervention to reduce HIV risk behaviors among drug users and risk partners in Chiang Mai, Thailand and Philadelphia, USA. Social Science and Medicine, 68(4), 740-748.

Lesko, C. R., Edwards, J. K., Cole, S. R., Moore, R. D., & Lau, B. (2017). When to censor?.

American Journal of Epidemiology, 187(3), 623-632.

VanderWeele, T. (2015). Explanation in causal inference: Methods for mediation and interaction. Oxford University Press.

Minor Comments

1.   Abstract: Spillover is one word. In the objectives, it says "spill over". I would also suggest saying "spillover among network members". When describing the participants, describe where they were recruited from and the study time frame. In the Results, please add

confidence intervals for the odds ratios. In the last sentence in the Results, can you please clarify what you mean by making a distinction between having change agents participate in the intervention versus change agents transmitting new knowledge?

2.   Introduction:

a.   Page 3, line 30: Please change "safer sexual practice" to "safer sexual practices" and "injection drug use" to "reducing injection drug use".

b.   Page 3, line 34: change "low adherence will" to "lower adherence is known to often".

c.   Page 3, line 34: When the authors say "neither method is fully effective" are they referring to both behavioral and biological interventions?

d.   Page 3, line 47: When the authors describe positive effects of interventions, it would help to be more specific about the interventions reported in the citations.

e.   Page 3, line 49: Avoid sentences in passive tense. Please rewrite this first sentence in the last paragraph on this page.

f.   Page 4, line 6: Do you mean "change to the outcome of a non-participant" rather than "change to a non-participant"?

g.   Page 4, line 18: A better word choice that "early participants" might be "initially recruited". The authors could also consider introducing the idea of the "index participant" for each network/cluster. In the Change Agents study, the CAs are the index participants.

h.   Page 4, line 23: Please change "initial population" to "initial

study population" and "increases" to "can improve".

i.   Page 4, line 27: Can you be more specific about the characteristics among individuals that may strengthen or weaken spillover effects? Are you referring to characteristics of the individuals, network structure, or both?

j.   Page 4, line 30: "Hidden networks" is jargon. Perhaps say "difficult networks to ascertain"?

k.   Page 4, line 33: Please change "an intervention" to "an investigator-initiated intervention".

l.   Page 4, line 43: Please change "enroll in the intervention" to "receive the intervention themselves".

m.   Page 5, line 3: Please change "transfer" to "are shared within social networks".

3.   Methods

a.   The methods could be clearer if the design of the study is described following the temporal ordering in the study (e.g., first describe what happened at recruitment, randomization, baseline, then the steps during follow-up).

b.   Page 6, line 24: Please change "significance" to "statistical significance of homophily".

c.   Page 7, line 15: Please change "the effect merely receiving an intervention" to "the effect of receiving the intervention". What is meant by "merely"?

d.   Page 7, line 18: Could you please clarify what you mean in this sentence: "For instance, this could occur if all CAs begin with good knowledge of HIV, and the intervention empowers them to convey knowledge they already had to their NMs". In this case, does the intervention provide any benefit to CAs?

4.   Results
a.   Page 8: Instead of predicting follow-up, I suggest predicting loss to follow-up.

b.   Page 8: Please clarify if this was the only CA variable (private water for CA) that predicted follow-up for NMs. What could explain this association? Please add this explanation to the Discussion Section.

c.   Page 8: When describing associations, please give the direction and magnitude, rather than saying if it was associated or not.

d.   Page 8: Was there anything else interesting to highlight from the Cox models, beyond that the results were closer to the null?

e.   Page 8, line 47: Please keep the units for the confidence interval the same as the point estimate (e.g., 6%, 20%).

f.   Page 8, line 50: Please change "HIV knowledge changing" to "HIV knowledge".

g.   Page 8, line 50: I find the interpretation of the NDE difficult to follow. Please clarify this sentence with more context from the study.

5.   Discussion

a.   Page 9, line 37: Instead of saying that you have shown spillover, please say
"evidence for spillover".

b.   Page 9, line 40: Please change "much information" to "maximal correct information".

c.   Page 9, line 46: Please clarify what you mean by "the population".

d.   Page 10, line 12: Please clarify the rationale for why it would be better to ascertain employed participants.

e.   Page 10, line 21: Please clarify what you mean by "if any of these characteristics acted in a causal fashion". In the article you estimated certain causal effects, so which characteristics are you referring to here? Please be more specific.

f.   Page 10, line 38: Was the average knowledge 80% among CAs? Or did 80% have perfect knowledge? Please clarify.

g.   Page 10, line 39: Please write "there was less room for improvements in knowledge" instead of a knowledge increase was not possible.

h.   Page 10, line 49: Please justify why the intervention would only spread one degree, as opposed to two or more, either citing existing literature or describing the mechanism of informational spread in the study population.

i.   Page 11, line 10: How could you determine highly connected participants in the social network?

j.   Page 11, line 14: Do you mean empowered network members to also be CAs, and trained by the CAs? Please clarify this sentence.

k.   Page 12, line 4: Please define what you mean by "closeness".

l.   Page 12, line 19: Please clarify future next steps. How could you examine the exact mechanisms of spillover?

6.   Tables and Figures

a.   Table 1: In the footnote, please clarify the differences described in the table. Please also add details about the statistical tests for each P value (e.g., which test you are using and which groups you are comparing).

b.   Table 2: For the characteristic "Complete HIV knowledge" is the upper confidence limit 125? Please check that the figure is

| | correct and that there are enough counts in each cell to estimate the odds ratio and its confidence interval.
c.   Figure (page 20): Please add a title and legend to this figure.

7.   Other:

a.   In the Strengths and Limitations section, when describing that the recruitment method helped to access an at-risk population, please give an example of other study designs that may not work as well. I would also suggest tempering this sentence by writing "typically not easily accessible". The second bullet could be reworded to be clearer (e.g., rather than the trial spillover, do you mean intervention spillover?). In the third bullet, what do you mean by the outcome being biased? Do you mean measurement of the outcome or estimation of outcome rates is biased? I believe you mean "estimation of the outcome will be biased if…".

b.   The authors describe there are no plans to disseminate findings to participants. Would this be possible? What possible steps could be taken to reduce the burden due to the intervention on participants in this study? Would it be possible to incentivize the intervention or apply contingency management principles for future studies?

c.   In some places, the authors use "treatment" and in others, "intervention". I suggest using the term "intervention" throughout, reserving the word "treatment" for a medical based intervention.

d.   Data Sharing Statement: Please clarify if the study data analyzed in this paper could be made available, and if so, under what conditions. |
|---|---|

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name

Hank Green

Institution and Country

Indiana University School of Public Health

Comment: **How many network members change agents were asked to enroll?**

Response: The network members were identified by the change agents themselves; therefore, the study staff did not recruit network members directly and this information is not available. In this regard, the network members may not represent the broader group of network members that were approached for recruitment. This may result in an underestimation of high risk behavior among network members. We have added some information about this in the Methods section (pg 6).

Comment: **What the demographics of the change agents and their network members were?**

Response: We include demographics of CAs and NMs, such as age, sex, education, and employment, in Table 1. The variables in Table 1 are exhaustive of variables included in all subsequent models.

Comment: Should justify the demographic and contextual variables included in the model—such as having a source of water. Etc. & **all variables used should be justified in the literature review and should be defined in a separate section**.

Response: Thank you for this comment! This is a good point, and we have thoroughly added explanations of each variable included in Tables 1 and 2 (p7).

Comment: **Should know what factors were associated with loss to follow-up for change agents**.

Response: This can be found in Siril et al., 2017 (reference #31), which focuses on the correlates of loss to follow-up (LTFU) among the change agents. In general, loss-to-follow-up among the change agents was very low (<10%) over the study period. The strongest correlate of LTFU was Anti-retroviral therapy, which would not be a factor for the many NMs not living with HIV. We have added language about this to the Methods. (p 6)

Comment: **Whether network members are lost because their related change agents are lost.**

Response: We have added "CA LTFU" as a regression variable which can be seen in Table 2. It is not significant as a predictor of LTFU among the NMs. (p 11)

Comment: Present results starting with descriptive statistics, then bivariate statistics, then multivariate models, summarizes the key results and then presents discussion. **For me, this would also describe the study, describe the change agents, describe the network members including how many on average were recruited per change agent.**

Response: Thank you for helping us clarify our approach. We have made sure that the analysis flows in the order you suggest: descriptive (Table 1 columns 1-3) then bivariate (Table 1 column 4), then multivariate (Table 2). Then, although it is essentially a bivariate model, due to its complexity and focus as a key finding, we then show the results from the spill over analysis. Then we move on to the summaries and discussion.
We have also expanded our description of CA and NM recruitment in the Methods, particularly focusing on the number of recruited NMs per CA. (p 6)

Comment: Assess whether there was bias in recruiting related to characteristics of the change agents.

Response: Thank you for this comment; bias in recruitment is often an issue in the generalizability of a study. Here, however, we were not primarily interested in a random selection of the population; we are most interested in who would serve best as a Change Agent in their community, and who could best cause spillover of the HIV knowledge and reduction of risky behaviors. The CA population was therefore not expected to perfectly reflect the characteristics of the general PLH population in Dar es Salaam. Furthermore, for the same reasons we did not collect information on NMs approached but not recruited, we did not collect information on CAs who were not recruited, exacerbated by the fact that they were recruited from HIV treatment centers.

Comment: Assess whether any of the clusters of dyads were influential enough to require adjustments to the standard errors related to clustering.

Response: Thank you for the comment. We ran the models with and without clustering, and the results were similar. However, we felt that for conceptual purposes, clustering is the appropriate analytical technique. This is especially true because one of our variables is the difference in age of NM and CA. If we have two NMs for the same CA, even if the age of the two NMs are uncorrelated, their differences between their age and the age of their CA will be correlated, because the same CA age is subtracted from both. Adjusting for these clusters is therefore the conservative approach that will address any correlations present therein. We added some discussion of this to the Methods section. (p 8)

Reviewer 2

Review of "Network Spillover Effects and Follow-Up Correlates in a HIV Prevention Intervention in Tanzania"

Major Comments

Comment: The definition of spillover seems narrow, could be clarified and also needs a reference.

Response: Thank you for this clarification. We have added some discussion about the definition and other names for spillover effects in the introduction. This includes some disambiguation of other terms such as behavioral spillover which can mean something different. We have also mentioned some of the other work done on this space. We were unable to find evidence that spillover effects of HIV knowledge interventions prevent HIV transmission in Sub-Saharan Africa, and have therefore added some discussion advocating for this as future research. (p 4 for definition; p 14 for partial interference)

Comment: More details are needed to describe the stepped wedge design of the trial and the implications for the analysis of spillover effects.

Response: Thank you for this comment. We have added a paragraph defining the exposure of the NMs based on when they completed their interview. This nicely divides the NMs into an exposed group (N=381) and an unexposed group (N=329). Because many NMs did not complete all of their surveys, using just their first survey provides a clean analysis of the stepped-wedge trial design. (p 6)

Comment: More details are needed about the change agents.

Response: This is an important comment, and we have clarified who the change agents are. We have defined them here based on their potential to become CAs. (p 5)
In terms of recruitment, each CA could recruit up to 3 NMs, and recruited an average of 1.07. Therefore, most CAs only have one CA-NM dyad, but some have larger egocentric networks in the study – necessitating our GEE and control for shared variance among dyads sharing a CA. (p 6)

Comment: **Please describe how dyads were ascertained from study data in the Methods Section.
Did you collect any information to look at possible contamination from intervention to control clusters?**

Response: We have added some definition about the CA-NM dyads to the methods section. We did not collect information to look at possible contamination, largely due to ethical concerns. Given the stigma of HIV infection, we did not ask CAs if they knew anyone else in the study, which would have been necessary to examine possible contamination. Given the large size of Dar es Salaam, and the small size of CA egocentric networks, we believe contamination is unlikely, but certainly not impossible. We have added some language about this to the limitations section. (p 6 for methods; p. 14 for limitations)

Comment: **The claim that the overlapping egocentric neighborhoods did not impact the analysis assuming the neighborhoods were non-overlapping is a strong assertion. Perhaps provide some insights into why randomly removing dyads is a valid way to test this assumption? If this assumption, does not hold, how would that impact your results?**

Thank you for this comment. This analysis was not about potentially overlapping neighborhoods of multiple CAs, but the egocentric networks of individual CAs having more than one NM not being independent. Removing NMs randomly so each CA only had one NM will test this assumption. We have added a paragraph to the methods section to clarify this. (p 8)
 As stated in the previous comments, we unfortunately do not have a way to examine the effects of overlapping neighborhoods. Since recruitment was at the clinic level and not the neighborhood level we anticipate that this would be less of a problem, since the clinic site received more than 5,000 patients per year.

Comment: Why did the authors restrict to a dyad analysis, when methods for egocentric networks exist?

Repsonse: Thank you for asking for clarification for this. Because CAs recruited only 1.07 NMs on average, we felt a full egocentric analysis was not warranted. The mediation analysis we used also requires independent dyads, which does not scale up to an egocentric analysis. We believe our sensitivity analysis of randomly removing the small number of CA-NM dyads so each CA only had a single dyad suffices to show that this approach was sufficient.

Comment: For the instrument used to collect knowledge, what was the Chronbach's alpha or Omega in your study sample? When describing the other studies that used this instrument, please be sure to describe the study population for South Africa and in the original study. Is it valid to combine across questions of this instrument to obtain a single measure (complete knowledge vs. incomplete)? Please provide some justification for this.

Response: Thank you for asking for this additional information, as it will help compare our population to others. We have added the Chronbach's alpha of this study, and have lavished additional detail on the other study populations. Further, we have clarified that an additional reason to combine across the questions was the lack of normality in the outcome variable, making a linear regression less ideal. (p 7)

Comment: Is using an intercept only GEE an appropriate test of homophily? Instead, the authors could consider an exponential random graph model to assess homophily and include nodal homophily terms. I did not see the homophily results described the Results Section. Please be sure to include these in the Results Section.

Response: Yes, we believe this is an appropriate test of homophily, and that an ERGM is not required. Because we have only egocentric data, and no knowledge about the potential ties between NMs, there is no structure which would need to be accounted for in our analysis. A tie-independent model therefore addressed our questions adequately. The analysis of homophily can be seen in the first paragraph of the results section. (p 8)

Comment: For the logistic regression in the second aim, how did the authors account for correlation in the study?

Response: Thank you for catching this! We actually fit a GEE here as well to account for the correlation of NMs having the same CA, but did not indicate this in the Methods section. We have changed this accordingly. (p 8)

Comment: Please clarify what you mean by interval censoring of the network members. Do you mean you defined their drop out date as halfway between their last visit and their subsequent missed visit? Please see the paper by Lesko, et al. about when to censor.

Response: Thank you for this clarification. This is a very helpful reference for exactly this problem. We have changed our censoring time to 3 months after their last interview, per the Lesko reference. (p 8 – ref #37).

Comment: For the third aim, please define what you mean by "the putative causal mechanism".

Repsonse: We have removed this phrase, and been much more explicit about the two mechanisms we are testing. (p 8)

Comment: I suggest instead of "spillover analysis", the authors could use "evaluation of spillover effects".

Response: we have changed our language to reflect this. (p 8)

Comment: I would suggest the authors estimate the direct, indirect, composite and overall effects as described in Hudgens and Halloran (2008) for clustered data or more specifically for dyads in Vanderweele's textbook (pages 399 – 402).

Response: Thank you for this comment. We have updated this citation and analysis accordingly. We use Vanderweele's method for dyads, and now our reference reflects that. (p 8)

Comment: Did the authors adjust for possible confounders in their analysis of spillover?

Response: Thank you for this question. We did not adjust for possible confounders because this was done in a randomized trial, and our exposed and unexposed NMs were balanced across all variables examined.

Comment: Would it be possible to adjust for selection bias due to loss to follow-up in your estimators of spillover? If not, this limitation is well covered in the discussion.

Response: The best way we can figure is to perform imputation of the HIV knowledge of those who were LTFU and rerun the analysis. If the results were similar, it would indicate that there was not a strong effect of selection bias. However, this approach would assume that those LTFU were the same as those not-LTFU, which seems like a stronger assumption than the one we make in our present analysis.

Comment: Is the partial interference assumption made in your analysis? If so, are there any doubts about this assumption? If so, please be sure to include in the Discussion Section.

Response: Thank you for this question, we were not aware of the formal definition of this assumption. We do make this assumption, and believe it is reasonable due to the size of Dar es Salaam, and the small cluster sizes. We have added this to our Discussion. (p 14)

Minor Comments
1. Abstract: Spillover is one word. In the objectives, it says "spill over". I would also suggest saying "spillover among network members". When describing the participants, describe where they were recruited from and the study time frame. In the Results, please add confidence intervals for the odds ratios. In the last sentence in the Results, can you please clarify what you mean by making a distinction between having change agents participate in the intervention versus change agents transmitting new knowledge?

Response: Thank you for these clarifications to the abstract. We have done all these.

2. Introduction: a. Page 3, line 30: Please change "safer sexual practice" to "safer sexual practices" and "injection drug use" to "reducing injection drug use".

Response: Done.

b. Page 3, line 34: change "low adherence will" to "lower adherence is known to often".

Response: Done.

c. Page 3, line 34: When the authors say "neither method is fully effective" are they referring to both behavioral and biological interventions?

Response: Yes. We clarified this.

d. Page 3, line 47: When the authors describe positive effects of interventions, it would help to be more specific about the interventions reported in the citations.

Response: We added some detail about these studies.

e. Page 3, line 49: Avoid sentences in passive tense. Please rewrite this first sentence in the last paragraph on this page.

Response: Thank you for catching our use of the passive voice. We have made things active, particularly the highlighted sentence.

12

f. Page 4, line 6: Do you mean "change to the outcome of a non-participant" rather than "change to a non-participant"?

Response: Yes we did; thank you for catching our typo. We have changed this language accordingly.

g. Page 4, line 18: A better word choice that "early participants" might be "initially recruited". The authors could also consider introducing the idea of the "index participant" for each network/cluster. In the Change Agents study, the CAs are the index participants.

Response: We have changed early to initially-recruited. We have for now not introduced the idea of index participant precisely because the index participant and CA is the same, and we do not want to complicate terminology.

h. Page 4, line 23: Please change "initial population" to "initial study population" and "increases" to "can improve".

Response: We have done so.

i. Page 4, line 27: Can you be more specific about the characteristics among individuals that may strengthen or weaken spillover effects? Are you referring to characteristics of the individuals, network structure, or both?

Response: We have added a sentence to clarify this. Although both of those are true, here we mean characteristics only as we are not examining the network structure itself in this study.

j. Page 4, line 30: "Hidden networks" is jargon. Perhaps say "difficult networks to ascertain"?
Response: We have changed this.

k. Page 4, line 33: Please change "an intervention" to "an investigator-initiated intervention".

Response: We have changed this.

l. Page 4, line 43: Please change "enroll in the intervention" to "receive the intervention themselves".

Response: We have changed this.

m. Page 5, line 3: Please change "transfer" to "are shared within social networks".

Response: We have changed this.

3. Methods a. The methods could be clearer if the design of the study is described following the temporal ordering in the study (e.g., first describe what happened at recruitment, randomization, baseline, then the steps during follow-up).

Response: This is a very good point, and we have rearranged the Study population section accordingly. We have not highlighted the whole section so the individual edits can remain clear.

b. Page 6, line 24: Please change "significance" to "statistical significance of homophily".

Response: We have done so.

c. Page 7, line 15: Please change "the effect merely receiving an intervention" to "the effect of receiving the intervention". What is meant by "merely"?

Response: We have done so. "Merely" was meant to highlight the fact that their response to the intervention was not important to the Natural Direct Effect (NDE) only that they received the intervention.

d. Page 7, line 18: Could you please clarify what you mean in this sentence: "For instance, this could occur if all CAs begin with good knowledge of HIV, and the intervention empowers them to convey knowledge they already had to their NMs". In this case, does the intervention provide any benefit to CAs?

Response: We have done so. We believe that in this case, the intervention is still useful. As we now indicate in the methods, we use the phrase "Change agent" optimistically. Unlike other studies which pick CAs based on characteristics like high in-degree, we allow CAs here to self-select. The fact that they had high HIV knowledge to begin with, but their NMs did not indicates that they had not previously felt empowered to share this knowledge. The high Natural Indirect Effect (NIE) indicates that the intervention helped them to do so, validating our calling them "CAs". According to the parent study, a number of outcomes improved for CAs, including an increase in self-efficacy for safer sex and a reduction in HIV-related stigma among other factors (see reference #26).

4. Results a. Page 8: Instead of predicting follow-up, I suggest predicting loss to follow-up.

Response: In line with the general framing of this paper, we chose to model follow-up due to the positive connotation of this approach. It does not lead one to subconsciously blame or judge those LTFU, but to appreciate what may lead participants to continue their participation in the study.

b. Page 8: Please clarify if this was the only CA variable (private water for CA) that predicted follow-up for NMs. What could explain this association? Please add this explanation to the Discussion Section.

Response: We have clarified this in the results, and have added some explanation to the discussion.

c. Page 8: When describing associations, please give the direction and magnitude, rather than saying if it was associated or not.

Response: We have added the point estimate for the characteristics mentioned without describing the magnitude.

d. Page 8: Was there anything else interesting to highlight from the Cox models, beyond that the results were closer to the null?

Response: We do not believe so. It was mainly meant as a sensitivity analysis since it cannot account for the egocentric clustering, and so the results are anti-conservatively biased.

e. Page 8, line 47: Please keep the units for the confidence interval the same as the point estimate (e.g., 6%, 20%).

Response: We have done so.

f. Page 8, line 50: Please change "HIV knowledge changing" to "HIV knowledge".

Response: We have done so.

g. Page 8, line 50: I find the interpretation of the NDE difficult to follow. Please clarify this sentence with more context from the study.

Response: We have clarified this in the context of the study, and included an interpretation that is in the form of a mediation effect so people familiar with that context can also better-interpret this finding.

5. Discussion a. Page 9, line 37: Instead of saying that you have shown spillover, please say "evidence for spillover".

Response: We have done so.

b. Page 9, line 40: Please change "much information" to "maximal correct information".

Response: We have done so.

c. Page 9, line 46: Please clarify what you mean by "the population".

Response: We have clarified what we meant here, to indicate that our RDS/snowball sampling procedure allowed us to recruit a more diverse sample than had we just focused on PLH in HIV treatment clinics.

d. Page 10, line 12: Please clarify the rationale for why it would be better to ascertain employed participants.

Response: This is an important point. We did not mean to suggest that unemployed persons should be excluded from research. We instead meant to highlight that employment may conflict with participation in research. We have therefore edited this sentence to indicate that the schedules of employed persons should be considered when conducting research to maximize participation.

e. Page 10, line 21: Please clarify what you mean by "if any of these characteristics acted in a causal fashion". In the article you estimated certain causal effects, so which characteristics are you referring to here? Please be more specific.

Response: Thank you for this comment. On reflection, this sentence was unclear and not integral to our discussion. We have thus removed it.

f. Page 10, line 38: Was the average knowledge 80% among CAs? Or did 80% have perfect knowledge? Please clarify.

Response: Thank you for this, we have clarified that the average score at baseline was 80%.

g. Page 10, line 39: Please write "there was less room for improvements in knowledge" instead of a knowledge increase was not possible.

Response: We have done so.

h. Page 10, line 49: Please justify why the intervention would only spread one degree, as opposed to two or more, either citing existing literature or describing the mechanism of informational spread in the study population.

Response: We have done so, showing that NMs would need to both increase their HIV knowledge and become empowered, whereas the CAs only need to become empowered. We have lessened the certainty of this conclusion to show that we don't know what would happen in this case, as the study does not examine a population with low baseline complete HIV knowledge.

i. Page 11, line 10: How could you determine highly connected participants in the social network?

Response: We have added some detail here and a reference showing that bridging ties are important for Change Agents.

j. Page 11, line 14: Do you mean empowered network members to also be CAs, and trained by the CAs? Please clarify this sentence.

Response: This could occur in one of two ways. 1) CAs only pass information on to NMs. In this case, the NMs could potentially spread the information further on, but would not officially become CAs as defined by NAMWEZA. 2) CAs could additionally receive training in leading their own NAMWEZA sessions. In this case, NMs would become CAs, and would spread the information and NAMWEZA further on. In this study, only (1) is possible. However, if we added training to lead NAMWEZA to future iterations, (2) could be possible as well. We have edited accordingly.

k. Page 12, line 4: Please define what you mean by "closeness".

Response: We have changed this to "relationally close" and changed the wording of this sentence a bit to make it more clear.

l. Page 12, line 19: Please clarify future next steps. How could you examine the exact mechanisms of spillover?

Response: Thank you for this comment. We have edited the section accordingly with a specific design to get at the mechanisms.

6. Tables and Figures a. Table 1: In the footnote, please clarify the differences described in the table. Please also add details about the statistical tests for each P value (e.g., which test you are using and which groups you are comparing).

Response: We apologize, but it is unclear what is meant by clarifying the differences. We have added details about the statistical tests we use.

b. Table 2: For the characteristic "Complete HIV knowledge" is the upper confidence limit 125? Please check that the figure is correct and that there are enough counts in each cell to estimate the odds ratio and its confidence interval.

Response: Yes, this is correct. There are enough counts in each cell, but not an overabundance.

c. Figure (page 20): Please add a title and legend to this figure.

Response: The caption and title should be at the end of manuscript. We have expanded the caption of the figure to more-directly relate to the NDE and NIE studied here.

a. In the Strengths and Limitations section, when describing that the recruitment method helped to access an at-risk population, please give an example of other study designs that may not work as well. I would also suggest tempering this sentence by writing "typically not easily accessible". The second bullet could be reworded to be clearer (e.g., rather than the trial spillover, do you mean intervention spillover?). In the third bullet, what do you mean by the outcome being biased? Do you mean measurement of the outcome or estimation of outcome rates is biased? I believe you mean "estimation of the outcome will be biased if…".

Response: Thank you for all of these comments, we have changed the strengths and limitations accordingly.

b. The authors describe there are no plans to disseminate findings to participants. Would this be possible? What possible steps could be taken to reduce the burden due to the intervention on participants in this study? Would it be possible to incentivize the intervention or apply contingency management principles for future studies?

Response: This would be possible for a subset of the population. During the study, we lost contact with a large number of NMs and some CAs. Given that the study completed 4 years ago, it is unlikely that we would be able to reach many NMs or CAs at this time.
Some sort of tele-participation might reduce the burden of the intervention, although we do not know how feasible that would be.

We did incentivize the intervention in this study, but not to the point that it would have overridden other financial concerns. Higher incentives might have helped with this.

c. In some places, the authors use "treatment" and in others, "intervention". I suggest using the term "intervention" throughout, reserving the word "treatment" for a medical based intervention.
d. Data Sharing Statement: Please clarify if the study data analyzed in this paper could be made available, and if so, under what conditions.

Response: Thank you for this comment; we have changed "treatment" to "intervention" when referring to the intervention (some uses of "treatment" do refer to medical-based interventions).

VERSION 2 – REVIEW

REVIEWER	Hank Green Indiana University School of Public Health
REVIEW RETURNED	02-Feb-2020

GENERAL COMMENTS	Thank you for the chance to review this revision. I believe the authors have been responsive to previous critiques and I have only minor concerns remaining: 1. Overall: make sure that the formatting and line breaks are correct. There are many places where words are split within a line and not between lines (for example page 5 line 27 and 37) 2. Abstract line 22: indent 'results' to follow the format of the rest of the abstract sections 3. Abstract line 36: I think intervening 'with' instead of intervening 'in' is a better way to discuss this. We are not aware of how well-connected the CAs are since network structure was not explored. I wouldn't say that your results 'indicate'. I would say your results 'suggest' 4. Page 4, line 18: please clarify this bullet point. I'm not sure what you're trying to say. 5. Page 6 line 7: 54 studies seems a lot if you don't compare to the total number of studies. Please include the total number so that 54 is revealed to be a small number of studies. 6. Page 7, line 7, remove the phrase 'and if they recruited… multiple dyads.' You should say if they recruited more than one NM they formed a cluster of ca-nm dyads. A network suggests that you collected information about the connections among the NMs and you did not do that. 7. Page 7 line 19: I think you mean to say that the CAs were one possible source of information for NMs. You are not entirely sure whether the CAs DID provide information directly to the NMs or if they might have also found information in other ways. 8. Page 7 Line 43: you should also report the percentage loss to follow up of your NMs here, even if it is reported in greater detail in another report. 9. Page 11, Table 1: refer to the journal's style sheet as to the p-

value, but normally they are posted as <0.001 and that is sufficient.

10. Page 12 line 50: I wouldn't say that these three differences count as 'many' significant differences. I would just say 'we found some significant differences between CAs and NMs'. And I would say that the NM recruits were different in some ways, but I would not say 'very' different. This is particularly true since two of your outcomes are dichotomous (education and employment) and a small change in education could have led to a difference based on how you dichotomized it.

REVIEWER	Ashley Buchanan University of Rhode Island United States
REVIEW RETURNED	03-Feb-2020

GENERAL COMMENTS	General Comments
	The authors were very responsive in their revisions and have addressed many of these comments. I found the paper was much easier to read and I now better understand their study design and analysis. The authors clarified details on their study design, including the change agents and how their ascertained their network members, refined the definition of spillover, and improved some aspects of their statistical methods. My remaining concerns are the lack of clarity about the network structure, the rationale for not adjusting for confounders, and conflating spillover methods with mediation analysis. Please find a summary of major comments and detailed minor comments below.
	Specific Comments for Revision
	Major Comments
	In the Introduction, please define what you mean by diffusion through social networks. When describing characteristics of individuals that may be more receptive to change, I suggest an important factor to also consider is knowledge about how many other individuals in proximity are infected. Do the authors think it might be worth introducing respondent driven sampling as a targeted sampling technique?
	For the measure "Complete HIV Knowledge", even if this is not a validated measure of true HIV knowledge, are there papers that could be cited to support this definition for the outcome? Are there any approaches in the literature to normalize this score, so it could be analyzed as continuous?
	The authors approach for assessing homophily seems reasonable. Would it be possible to provide a citation for this approach? I am still not clear if the analysis included dyads only, or egocentric networks. Could the authors please make this point clear? Do the authors know how many CAs had multiple network members? Please clarify this in the methods section. In the response to earlier comments, can the authors explain what the difference is between overlapping neighborhoods and CAS having more than one NM not being independent? Those sound like the same concept.
	For the definition of loss-to-follow-up, according to the Lesko paper,

the authors should censor at the time of their last interview (not with the additional three months) because the outcome was ascertained from within the study, not an external data source. If a CA was lost to follow-up, were their NMs also lost to follow-up? The authors should also make clear their assumption about the missing data, which I believe is missing completely at random.

1

Even though this is a randomized trial, the authors need to consider if the exposures of interest were randomized. When making comparisons between outcomes of CAs and NMs, the CA/NM status was not randomized, even though the study design was step-wedge randomized. My understanding is that the CAs were also not randomized either to be educators or not. Is that the case? If so, do the authors need to control for individual-level confounders? What about temporal confounding factors?

The authors still seem to conflate mediation with spillover. The natural indirect effect and the natural direct effect are terminology from the meditation literature. Spillover is the effect on the CA's exposure on the other dyad member's HIV knowledge. This is a distinct concept from an exposure affecting an outcome through a mediator. In this paper, there does not seem to be a mediator, but rather only interest in estimating spillover effects. I would suggest the authors estimate the direct, indirect, composite and overall effects as described in Hudgens and Halloran (2008) for clustered data or more specifically for dyads in Vanderweele's textbook (pages 399 – 402). Do the authors have a reference for the SAS macro used, in addition to a published paper that applies these methods?

For the modeling of follow-up of NMs, why did the authors use odds ratios, instead of risk/rate ratios? I believe this analysis is longitudinal (i.e., not cross sectional or a case-control study), so risk ratios would be more appropriate. These can be estimated using a log-binomial regression (Spiegelman, 2005). I also recommend the authors include a mention of the limitations of hazard ratios in the discussion section (Hernan, 2010). Hazard ratios have built-in selection bias and provide only a single estimate across all time points. The recommendation is to instead provide survival curves and risk difference/ratio measures at specific time points of interest. Because the hazard ratios are from a sensitivity analysis, the parameters are fine as reported, but a discussion of this point would be helpful.

How is the claim that the intervention effects will only spillover by one degree if your sample does not include two-degree contacts or more? Please soften this claim or justify with your results.

For the partial interference assumption, the authors should introduce this assumption and a more complete discussion of important assumptions in the Methods section. In the Discussion section (page

15, line 7), I believe the authors meant to say contamination would lead to a violation of the partial interference assumption. I also suggest discussing contamination and a possible violation of the partial interference assumption separately.

References:

Hernán, M. A. (2010). The hazards of hazard ratios. Epidemiology (Cambridge, Mass.), 21(1), 13.

Lesko, C. R., Edwards, J. K., Cole, S. R., Moore, R. D., & Lau, B. (2017). When to censor?.

American Journal of Epidemiology, 187(3), 623-632.

2

Spiegelman, D., & Hertzmark, E. (2005). Easy SAS calculations for risk or prevalence ratios and differences. American Journal of Epidemiology, 162(3), 199-200.

VanderWeele, T. (2015). Explanation in causal inference: Methods for mediation and interaction. Oxford University Press.

Minor Comments

1.   Introduction:

a.   Page 3, line 40: Please explain the 90-90-90 treatment goals. Or simply describe more broadly.

b.   Page 5, line 5: Better word choice for "HIV-risky behaviors"? "HIV risk behaviors".

c.   Page 5, line 18: Reword "interventions have shown spillover effects…" to "HIV prevention educational interventions were demonstrated to have spillover effects…"

d.   Page 5, line 22: Please clarify what the intervention is and what the outcome is in this sentence. I believe you mean spillover effects for the HIV knowledge intervention.

e.   Page 5, line 30: Clarify what you mean by "directly examined". Same comment for "directly educated". Someone is educated or not.

2.   Methods

a.   Page 6, line 15: Instead of "Interventions lasted…", the authors could change this to "follow-up lasted". Was the intervention continuously delivered for the entire follow-up? Or do the authors mean the entire study duration was from 2010 to 2014? Please clarify.

b.   Page 6, line 18: Do the authors mean that the study staff did

not deliver the intervention to the participants? Please clarify this sentence.

c.   Page 6, line 28: Please clarify what you mean by "longitudinal spillover". This could be made clearer by explaining the structure of the follow-up data first. Do you have pre and post measurements on both the NMs and CAs?

d.   Page 6, line 42: Please report the proportion loss-to-follow-up among the NMs.

e.   Page 8, line 22: The writing in this paragraph could be clarified to explain what the outcome was. In some places, it is described as "follow-up" and in others "loss to follow-up". I would suggest explaining what the event indicator was, which might be "completed follow-up" and the time is measured from randomization to the three months after their last clinic visit.

f.   Page 8, line 42: Perhaps instead of saying "does not result in biased estimates" say "so in this case, the method is appropriate".

g.   This is very minor, but in some places the authors write "dyad" and "pair" in the other. Perhaps use one of these words? Or make clear that these have the same meaning.

h.   Are the sensitivity results for removing multiple NMs for a single CA described somewhere?

3

3.   Results

a.   Does this result even need a statistical test: "Only 12.3% of NMs were HIV-positive, compared to all CAs"?

b.   For the odds ratio for "having complete HIV knowledge", are the authors sure there is not a positivity violation for this variable (i.e., small expected cell counts)?

c.   Page 10, line 28: Do the authors mean "completed follow-up" instead of "followed-up"?

d.   Please add confidence intervals for the hazard ratios reported in the results section.

4.   Discussion

a.   Page 11, line 46: Maximal is used twice in this sentence. Perhaps a different word choice?

b.   Page 12, line 13: Please link the sentence about the Latkin (2013) study to the prior sentence. I do not see the connection as written.

c.   Adjusting for selection bias using censoring weights might help this issue. The authors could mention this in the discussion.

	d. Page 13, line 25: Please clarify what the authors mean by "the effects seen here could spillover continuously".
	e. Page 14, line 20: What do the authors mean by "relationally close"? Please clarify.
	f. Page 14, line 42: Please use more scientific language to replace "putting individuals at risk…".
	g. I did not find any discussion about the association between private water for CA predicting follow-up for NM. Could the authors please point me to this? Or add this if the authors have not already done so?
	5. Tables and Figures
	a. In Table 2, please clarify what the outcome was for the Cox PH model. The time is follow-up time. What is your event indicator?
	b. In Table 1, please clarify what you mean by "difference for continuous variables" or "concordance.."? What is the outcome and exposure in these models?
	6. Other:
	a. In Strengths and Limitations, please change "factors leading to drop out" to "factors associated with drop out".
	b. Data Sharing Statement: Please clarify if the study data analyzed in this paper could be made available, and if so, under what conditions.

VERSION 2 – AUTHOR RESPONSE

<u>**Reviewer 1**</u>

1. **Overall**: make sure that the formatting and line breaks are correct. There are many places where words are split within a line and not between lines (for example page 5 line 27 and 37)

   Response: Thank you for this comment. We did miss a number of these and have gone through to correct them. (throughout)

2. **Abstract** line 22: indent 'results' to follow the format of the rest of the abstract sections

   Response: We have done so. (page 2)

3. **Abstract** line 36: I think intervening 'with' instead of intervening 'in' is a better way to discuss this. We are not aware of how well-connected the CAs are since network structure was not explored. I wouldn't say that your results 'indicate'. I would say your results 'suggest'

Response: Thank you for identifying these wording issues. You are correct that we did not fully-explore the networks of CAs, so we have moderated out language to 'suggesting'. We have also changed "with" to "among". (page 2)

4. **Page** 4, line 18: please clarify this bullet point. I'm not sure what you're trying to say.

   Response: We have re-written this bullet point to indicate that the high dropout was simultaneously a strength and a limitation, and we also explain why this is the case. (page 3)

5. **Page** 6 line 7: 54 studies seems a lot if you don't compare to the total number of studies. Please include the total number so that 54 is revealed to be a small number of studies.

   Response: Thank you for making this point. The cited review does not include a number of spillover studies overall. However, in Supplement 7, Figure 1, it shows the flow diagram of studies included and excluded. A total of 736 studies (of 12,784 abstracts) had their full text screened, indicating that the abstract had sufficient evidence of spillover as a metric of interest to merit full-text screening. The number 736 is therefore what we include in the paper (page 5).

6. **Page** 7, line 7, remove the phrase 'and if they recruited… multiple dyads.' You should say if they recruited more than one NM they formed a cluster of ca-nm dyads. A network suggests that you collected information about the connections among the NMs and you did not do that.

   Response: Thank you for requesting this change. You are correct that we do not measure egocentric networks, so we have changed this as indicated (page 6)

7. **Page** 7 line 19: I think you mean to say that the CAs were one possible source of information for NMs. You are not entirely sure whether the CAs DID provide information directly to the NMs or if they might have also found information in other ways.

   Response: Thank you for this clarification. This is correct; the NM could have received information from a variety of sources. The differential increase in knowledge between the intervention and control indicates that it is *likely* due to the CA undergoing NAMWEZA, but of course not guaranteed. We have removed this clause, and believe the paragraph now reads better and more accurately (page 6).

8. **Page** 7 Line 43: you should also report the percentage loss to follow up of your NMs here, even if it is reported in greater detail in another report.

   Response: We have done so. (Page 6).

9. **Page** 11, Table 1: refer to the journal's style sheet as to the p-value, but normally they are **posted** as <0.001 and that is sufficient.

   Response: Thank you for pointing this out. We have changed p-values <0.0001 to <0.001. (Throughout)

10. **Page** 12 line 50: I wouldn't say that these three differences count as 'many' significant differences. I would just say 'we found some significant differences between CAs and NMs'.

And I would say that the NM recruits were different in some ways, but I would not say 'very' different. This is particularly true since two of your outcomes are dichotomous (education and employment) and a small change in education could have led to a difference based on how you dichotomized it.

Response: Thank you this comment; we have changed the text as suggested. (page 12)

Reviewer 2

Comment: In the Introduction, please define what you mean by diffusion through social networks. When describing characteristics of individuals that may be more receptive to change, I suggest an important factor to also consider is knowledge about how many other individuals in proximity are infected. Do the authors think it might be worth introducing respondent driven sampling as a targeted sampling technique?

Response: Thank you for this comment. We have defined what we mean by diffusion in this context. Although we agree that an important factor in diffusion is knowledge about how many other individuals may be affected, it is not as relevant in the context of this manuscript, as we cannot directly examine it. We do not want to lead our readers to unmet expectations by discussing "complex contagion" (multiple infected neighbors are required for one to become infected) in the Introduction.

With respect to RDS, our sampling scheme was effectively RDS going a single degree away from the initial participant, although we do not call it such. Reference [23] includes a discussion of RDS, and we have changed the introduction to explicitly mention it. (page 5)

Comment: For the measure "Complete HIV Knowledge", even if this is not a validated measure of true HIV knowledge, are there papers that could be cited to support this definition for the outcome? Are there any approaches in the literature to normalize this score, so it could be analyzed as continuous?

Response: Thank you for this comment; given that this is not validated and not standard practice, we should justify this choice. We have added two citations to the paper to do so. The first shows an analysis where HIV knowledge was coded as a series of dichotomous variables; each one was therefore complete/incomplete on a single question. We merely extend this to a total of seven questions. The second shows that in general, dichotomization is often not appropriate and reduces power. However, our analysis is a case where it is appropriate, as our data are highly-skewed. These arguments combined support our approach of using dichotomization. (page 7)

Comment: The authors approach for assessing homophily seems reasonable. Would it be possible to provide a citation for this approach?

Response: Thank you for your comment. We chose this approach as a parametric alternate to a bootstrapped/permuted data tests for homophily. However, we were unable to find another paper doing this, and so we have changed our method to a standard permutation test to make the method more generally-acceptable (reference 39). No significance levels changed in this alternative approach. (page 8)

Comment: I am still not clear if the analysis included dyads only, or egocentric networks. Could the authors please make this point clear?

Response: Thank you for this comment; we have modified our language slightly, as we do not have complete egocentric networks. Rather the analysis is on a set of CA-NM dyads, some of which share a common CA. (top of page 6)

Comment: Do the authors know how many CAs had multiple network members? Please clarify this in the methods section.

Response: Yes we do, and we have added a clause about this in the Methods section. (bottom of page 6)

Comment: In the response to earlier comments, can the authors explain what the difference is between overlapping neighborhoods and CAS having more than one NM not being independent? Those sound like the same concept.

Response: Thank you for requesting this clarification. We have added some explanation of NMs not being independent in the first paragraph of the Statistical Analysis subsection. (page 7-8)

Comment: For the definition of loss-to-follow-up, according to the Lesko paper, the authors should censor at the time of their last interview (not with the additional three months) because the outcome was ascertained from within the study, not an external data source.
Response: Thank you for this comment; we misinterpreted the Lesko paper. We now correctly use the last study encounter as the time of lost-to-follow-up. (page 8)

Comment: If a CA was lost to follow-up, were their NMs also lost to follow-up?
Response: Not by design/definition. As can be seen in Table 2, if the CA was lost to follow-up, it was associated with a non-significant increase in the NM's likelihood of not completing a follow-up interview.

Comment: The authors should also make clear their assumption about the missing data, which I believe is missing completely at random.
Response: Thank you for this comment; we have specified our assumption of missingness, which is indeed MCAR. (page 6)

Comment: Even though this is a randomized trial, the authors need to consider if the exposures of interest were randomized. When making comparisons between outcomes of CAs and NMs, the CA/NM status was not randomized, even though the study design was step-wedge randomized. My understanding is that the CAs were also not randomized either to be educators or not. Is that the case? If so, do the authors need to control for individual-level confounders? What about temporal confounding factors?
Response: Thank you for asking for this clarification. The status of CA/NM was not randomized, so we do not expect Cas and NMs to be identical. We have added language to this effect in the Methods. (page 7-8)
What was randomized is *when* the CA received *NAMWEZA*. The stepped-wedge design ensured that all CAs were exposed to *NAMWEZA*, and thus became educators, at some point.
The initial study design proposed doing 3 surveys of the NMs – one after each wave. Under that paradigm, we could then do a more-complex analysis of the NMs over the course of the trial, treating them as exposed for all interviews after their CA was exposed to *NAMWEZA*. However, most NMs only completed a baseline and one follow-up interview (mostly after the first wave), so we instead did a pre-post analysis treating NMs as exposed if their CA had participated in *NAMWEZA* prior to the NMs follow-up interview. So for the overall study, we found that NMs whose CA had randomly been assigned *NAMWEZA* in the first wave increased their proportion of complete HIV knowledge by 12% more than the NMs whose CA was not randomly assigned *NAMWEZA* in the first wave (was instead assigned in the second or third wave). In this analysis, NM exposure was randomized, and so we do not need to control for confounders. We have added some language to the Methods to further explain this. (bottom of page 8)

Comment: The authors still seem to conflate mediation with spillover. The natural indirect effect and the natural direct effect are terminology from the meditation literature. Spillover is the effect on the CA's exposure on the other dyad member's HIV knowledge. This is a distinct concept from an exposure affecting an outcome through a mediator. In this paper, there does not seem to be a mediator, but rather only interest in estimating spillover effects. I would suggest the authors estimate the direct, indirect, composite and overall effects as described in Hudgens and Halloran (2008) for clustered data or more specifically for dyads in Vanderweele's textbook (pages 399 – 402). Do the authors have a reference for the SAS macro used, in addition to a published paper that applies these methods?
Response: We thank the reviewer for this point. We realize we made some of our analytic approach and justification unclear. In editing the Methods for the point above, we have also added language indicating that the setup of our study allowed us to examine spillover effects by way of a mediation analysis. Here, a CA being exposed to NAMWEZA is the exposure, a CA's change in HIV knowledge is the potential mediating variable, and the outcome is the NM's change in HIV knowledge (Figure 1). To evaluate a spillover via mediation in this way requires a number of assumptions, which are likely met due to the study design.
We have added a reference directly leading to the SAS macro we used and a paper applying the method. (reference 48, page 9)

Comment: For the modeling of follow-up of NMs, why did the authors use odds ratios, instead of risk/rate ratios? I believe this analysis is longitudinal (i.e., not cross sectional or a case-control study), so risk ratios would be more appropriate. These can be estimated using a log-binomial regression (Spiegelman, 2005).

Response: We thank the reviewer for this point; presenting RRs is indeed appropriate here, and we have changed the model to a log-binomial model. The results generally remain the same direction and significance as a result. (page 8 in Methods, page 11 for Results)

Comment: I also recommend the authors include a mention of the limitations of hazard ratios in the discussion section (Hernan, 2010). Hazard ratios have built-in selection bias and provide only a single estimate across all time points. The recommendation is to instead provide survival curves and risk difference/ratio measures at specific time points of interest. Because the hazard ratios are from a sensitivity analysis, the parameters are fine as reported, but a discussion of this point would be helpful.

Response: Thank you for this comment. We have added a mention of the limitations of hazard ratios. (page 8)

Comment: How is the claim that the intervention effects will only spillover by one degree if your sample does not include two-degree contacts or more? Please soften this claim or justify with your results.

Response: We thank the reviewer for pointing this out. We believe there was a misunderstanding in our language. We mean that spillover will only occur one degree in the sense of "six degrees of separation", that is, A->B->C is two degrees if A is not directly connected to C. So, if one CA is directly connected to multiple NMs (has a degree > 1), there would be spillover to them, as they are all one degree away from A (one degree of separation). We have modified the use of degree in this context to ensure clarity. (page 13)

Comment: For the partial interference assumption, the authors should introduce this assumption and a more complete discussion of important assumptions in the Methods section. In the Discussion section (page 15, line 7), I believe the authors meant to say contamination would lead to a violation of the partial interference assumption. I also suggest discussing contamination and a possible violation of the partial interference assumption separately.

Response: Thank you for this comment. We have added some discussion of this to the Methods section, including some discussion of contamination. We have removed that limitation from the Discussion section, which is now discussed in the Methods. (page 9)

1. Introduction:
a. **Page** 3, line 40: Please explain the 90-90-90 treatment goals. Or simply describe more broadly.
Response: We have added this.

b. **Page** 5, line 5: Better word choice for "HIV-risky behaviors"? "HIV risk behaviors".
Response: We have changed this accordingly.

c. **Page** 5, line 18: Reword "interventions have shown spillover effects…" to "HIV prevention educational interventions were demonstrated to have spillover effects…"
Response: We have changed this accordingly.

d. **Page** 5, line 22: Please clarify what the intervention is and what the outcome is in this sentence. I believe you mean spillover effects for the HIV knowledge intervention.
Response: Thank you for this comment. We have edited this accordingly, and hope it also helps to clarify our previous response about conflating mediation and spillover.

e. **Page** 5, line 30: Clarify what you mean by "directly examined". Same comment for "directly educated". Someone is educated or not.
Response: We have removed "directly" from "directly examined" and changed "directly educated" to "educated … one-on-one".

2. Methods

a. **Page** 6, line 15: Instead of "Interventions lasted…", the authors could change this to "follow-up lasted". Was the intervention continuously delivered for the entire follow-up? Or do the authors mean the entire study duration was from 2010 to 2014? Please clarify.
Response: We have changed this statement to more-clearly indicate that we mean the study duration was from 2010 to 2014 across all waves, and that the last interview (and therefore followup) went through March 2014.

b. **Page** 6, line 18: Do the authors mean that the study staff did not deliver the intervention to the participants? Please clarify this sentence.
Response: Yes, this is what we meant. We have changed this accordingly.

c. **Page** 6, line 28: Please clarify what you mean by "longitudinal spillover". This could be made clearer by explaining the structure of the follow-up data first. Do you have pre and post measurements on both the NMs and CAs?
Response: Thank you for this comment. We have removed the word "longitudinal" from this, and have clarified that we have baseline and follow-up surveys for both CAs and NMs, and for some NMs, this occurs before the CA had the intervention, but for others, it was after their CA had the intervention.

d. **Page** 6, line 42: Please report the proportion loss-to-follow-up among the NMs.
Response: We have done so.

e. **Page** 8, line 22: The writing in this paragraph could be clarified to explain what the outcome was. In some places, it is described as "follow-up" and in others "loss to follow-up". I would suggest explaining what the event indicator was, which might be "completed follow-up" and the time is measured from randomization to the three months after their last clinic visit.
Response: We have clarified the language here to make it clear that both models are using the same outcome: "completed follow-up" or equivalently "completed a follow-up interview".
NB: We have also changed our censoring time per the other comment, and participants are now censored at the time of their latest interview, rather than 3 months later.

f. **Page** 8, line 42: Perhaps instead of saying "does not result in biased estimates" say "so in this case, the method is appropriate".
Response: We have done so.

g. **This** is very minor, but in some places the authors write "dyad" and "pair" in the other. Perhaps use one of these words? Or make clear that these have the same meaning.
Response: We have changed uses of "pair" to "dyad".

h. **Are** the sensitivity results for removing multiple NMs for a single CA described somewhere?
Response: Thank you for catching this. No, they are not. We have added a sentence saying the results were very similar when performing this analysis.

3. **Results**
a. **Does** this result even need a statistical test: "Only 12.3% of NMs were HIV-positive, compared to all CAs"?
Response: We believe it does. Although CAs were asked to recruit NMs they believed to be at-risk of HIV, it was not guaranteed that they would do so. The statistical test formalizes the success of the recruitment strategy.

b. **For** the odds ratio for "having complete HIV knowledge", are the authors sure there is not a positivity violation for this variable (i.e., small expected cell counts)?
Response: Yes, although the rates of complete HIV knowledge were high in both CAs and NMs, the minimum cell size was 64, well above the size at which a non-parametric test would be preferred.

c. **Page** 10, line 28: Do the authors mean "completed follow-up" instead of "followed-up"?
Response: We have changed this to "completed a follow-up interview".

d. **Please** add confidence intervals for the hazard ratios reported in the results section.
Response: We have done so. We have also added hazard ratios to some of the ORs in the preceding paragraph for the sake of completeness.

4. Discussion
a. **Page** 11, line 46: Maximal is used twice in this sentence. Perhaps a different word choice?
Response: We have changed the second use of "maximal" to "ensure study information".

b. **Page** 12, line 13: Please link the sentence about the Latkin (2013) study to the prior sentence. I do not see the connection as written.
Response: We have added a clause about this.

c. **Adjusting** for selection bias using censoring weights might help this issue. The authors could mention this in the discussion.
Response: We have mentioned this in the limitations section.

d. **Page** 13, line 25: Please clarify what the authors mean by "the effects seen here could spillover continuously".
Response: We have added a second layer of what would happen in this scenario to lead to self-perpetuating interventions.

e. **Page** 14, line 20: What do the authors mean by "relationally close"? Please clarify.
Response: We have changed this wording somewhat to be more about the mental heuristics CAs used to decide who to nominate. Further understanding how they decided how to nominate and recruit NMs may lead to a better understanding of how to elicit the types of NMs a study ideally wants.

f. **Page** 14, line 42: Please use more scientific language to replace "putting individuals at risk…".
Response: We have changed this to "behaviors which may increase one's risk of contracting HIV.

g**. I did not** find any discussion about the association between private water for CA predicting follow-up for NM. Could the authors please point me to this? Or add this if the authors have not already done so?
Response: Thank you for this observation. We did not include an explicit discussion of this association, and have now added it in.

5. Tables and Figures
a. **In Table 2**, please clarify what the outcome was for the Cox PH model. The time is follow-up time. What is your event indicator?
Response: We have added some information to this effect in the figure caption.

b. **In Table 1**, please clarify what you mean by "difference for continuous variables" or "concordance.."? What is the outcome and exposure in these models?
Response: We have added examples of the difference and concordance calculation. In these models, there was not an exposure; we use the intercept-only model as we are interested in whether the difference was non-zero, or whether concordance was not 50%; the GEE framework allowed us to control for repeated CA measures.

6. Other:
a. **In Strengths and Limitations**, please change "factors leading to drop out" to "factors associated with drop out".
Response: We have done so.

b. **Data Sharing Statement**: Please clarify if the study data analyzed in this paper could be made available, and if so, under what conditions.
Response: We have done so.

Editorial requests:

Comment: Please revise the title of your manuscript to include the research question, study design and setting. This is the preferred format of the journal.
Response: We have changed the title to "Evaluating spillover of HIV knowledge from study participants to their network members in a stepped-wedge behavioral intervention in Tanzania"

Comment: Please ensure that all abbreviations are defined on first mention, including those in the abstract.
Response: We have done so, with the exception of HIV and AIDS, which we believe is common knowledge.

Comment: Please clarify whether informed verbal or written consent was obtained from all participants in your study.
Response: We have done so.

Comment: Please ensure that all citations in your manuscript are correctly referenced (e.g. we note that there is an error in your citation on page 6).
Response: We have done so.

VERSION 3 – REVIEW

REVIEWER	Ashley Buchanan University of Rhode Island, USA
REVIEW RETURNED	12-Mar-2020

GENERAL COMMENTS	General Comments
	The authors were very responsive in their revisions and have addressed many of these comments. I found the writing greatly improved and the article is much easier to follow. The authors clarified their statistical methods, study design, and limitations. My remaining concerns are the appropriateness of mediation analysis to evaluate spillover and clarification of the assumptions in the analysis. Please find a summary of major comments and detailed minor comments below. Specific Comments for Revision Major Comments My concern remains that the authors are conflating spillover with mediation analysis. Please see the footnote page 401 in Vanderweele's book, the reference cited for the authors use of this method. In the second paragraph on that page, he writes "this is very different from the indirect effects we were considering in Part I of this book in the context of mediation in which an exposure affected an outcome through a mediator". Perhaps one can frame spillover as a mediation analysis, but the authors need to provide a stronger rationale for this, or ideally a citation that explains this method. Vanderweele's book does not support using mediation analysis to evaluate spillover. Alternatively, the authors could review Chapter 15 and apply appropriate models to analyze spillover effects. The methods section could benefit from more details about the timing of the intervention in the stepped-wedge RCT. The authors should explain the three waves more clearly, as well as a sentence about the stepped-wedge design. The authors now clearly explain

the timing of the intervention to the CA and when the outcomes were assessed for NMs.

Education exposure was randomized by cluster (to the CA and their network members), but the role of each person not randomized CA/NM. Does that impact the analysis for any of the aims? Do the authors compare outcomes between CA and NM? If this does not matter, could the authors please explain this in the text?

The authors describe the loss-to-follow-up. Was there any other information missing at baseline or follow-up (exposure, outcome, covariates)? If so, please describe. Perhaps the sentence about loss-to-follow-up could be moved closer to the discussion of the MCAR assumption.

For the permutation test, what are the assumptions to perform this test? Are they met in the analysis? Please clarify for the reader. Can the authors be more specific about the statistics they computed in the 1000 permuted samples?

Did the authors use one NM per CA for the mediation analysis? Or multiple NM per CAs? The sentence at the bottom of page 11 sounds like multiple NMs were used in the analysis. Does this violate any assumptions of that approach?

Minor Comments

1.   Introduction:

a.   Page 7, line 13: Do the authors mean to increase the reach of the intervention, rather than enrollment?

2.   Methods

a.   Page 7, line 31: What do the authors mean by "foster through the intervention? Please clarify.

b.   Page 8, line 27: What do the authors mean by "net of temporal or geographical trends"? Please clarify. Perhaps this is a point for the statistical methods section that relates to the methods used?

c.   Were all demographic and contextual variables measured at baseline and prior to the exposure? Please add that point. For the effects estimated, it is important to have the intervention prior to the outcome and the covariates measured at baseline (prior to the intervention).

d.   Page 10, line 29: I would suggest saving the discussion of the limitations of hazard ratios for the discussion section. The second limitation is that it provides only one estimate for the duration of the study.

e.   Should the assumptions be reported before the methods? What other assumptions are made in the three parts of the analysis? The second aim treats the network members as independent. Is the third aim assuming that network members and CA are exchangeable, perhaps conditional on covariates?

	f. Page 10, line 40: The authors should clarify that these models are about loss to follow-up for the NM. 3. Results a. Does this result even need a statistical test: "Only 12.3% of NMs were HIV-positive, compared to all CAs"? How did the authors perform a statistical test if one of the cells was zero? 4. Discussion a. Page 13, line 47: Please clarify for which analyses the loss to follow-up may have biased. This bias is particularly a concern if the dropout was different across intervention arms and associated with the outcome. b. Where is the discussion about contamination and a possible violation of the partial interference assumption? 5. Tables and Figures a. In the figure on page 23, what are the different arrows (dash vs. solid)? References: VanderWeele, T. (2015). Explanation in causal inference: methods for mediation and interaction. Oxford University Press.

VERSION 3 – AUTHOR RESPONSE

Major comments:

Comment: My concern remains that the authors are conflating spillover with mediation analysis. Please see the footnote page 401 in Vanderweele's book, the reference cited for the authors use of this method. In the second paragraph on that page, he writes "this is very different from the indirect effects we were considering in Part I of this book in the context of mediation in which an exposure affected an outcome through a mediator". Perhaps one can frame spillover as a mediation analysis, but the authors need to provide a stronger rationale for this, or ideally a citation that explains this method. Vanderweele's book does not support using mediation analysis to evaluate spillover. Alternatively, the authors could review Chapter 15 and apply appropriate models to analyze spillover effects.

Response: We thank the reviewer for bringing this up again. The fact that it remains unclear indicates we did not explain our position adequately. We remain convinced that in this paper, we are in fact evaluating spillover via a causal mediation analysis, i.e., whether the effect of one person's exposure on another person's outcome is mediated by the first person's outcome or not.

From pages 402-403 of VanderWeele's 2015 textbook: "The methods described below could also be employed in a study in which the exposure were, say, a smoking cessation program in which one of two persons in a household participated. The participation of the first person might affect the smoking behavior of the second. This might occur either (i) because smoking cessation for the first person encourages the second to stop smoking or (ii) because even if the first person does not stop smoking, the second person might nevertheless be exposed to some of the smoking cessation program materials. One could potentially assess the presence of this second type of effect by applying the methods described below concerning the "infectiousness effect."

This example is exactly analogous to our study. Instead of smoking cessation, the behavior in question is risky sexual behavior. 1) is the direct effect where the intervention causes a CA to increase their HIV knowledge and pass that knowledge onto the NM 2) is the indirect effect, where the intervention may not directly increase the CA's knowledge, but may still mention being a CA etc., leading to the NM seeking information on their own.

A more recent study also using this approach to study spillover, and the authors wrote the following: "In this paper, we consider the causal mechanisms of this average spillover effect by using the idea from the causal mediation literature (e.g. Robins and Greenland (1992), Pearl (2001), Imai et al . (2010) and VanderWeele (2015)). Specifically, we decompose the average spillover effect into the sum of the contagion effect and direct effect by considering the vote intention of the contacted voter as the mediator." We have added this reference to the paper and elaborated some of the above in the paper as well.

Relevant Text: These pathways represent different types of spillover effects: the exposure or outcome of one person affecting the outcome of another person.
As shown by VanderWeele et al (2015), social network spillover effects in the case of dyadic relationships can be broken down into concepts from mediation analysis: direct and indirect effects (Figure 1) [42]. This method has since been used for novel evaluations of spillover effects [43,44].
[44] Imai K, Jiang Z. Identification and sensitivity analysis of contagion effects in randomized placebo-controlled trials. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2019.

Comment: The methods section could benefit from more details about the timing of the intervention in the stepped-wedge RCT. The authors should explain the three waves more clearly, as well as a sentence about the stepped-wedge design. The authors now clearly explain the timing of the intervention to the CA and when the outcomes were assessed for NMs.

Response: Thank you for this comment. An additional reading of this section did indeed lack detail on the stepped-wedge design and the waves. We have added two sentences to better-describe this. We also added a sentence directing inquiring readers to the study design paper which further lays out the details.

Relevant Text: As fits a stepped-wedge RCT, all CAs eventually received the intervention, but were randomized to when they received it. These waves each lasted 12 weeks, at which point the next wave began and another group of CAs received the intervention.
For more information on the study design, we direct interested readers to Smith-Fawzi et al., 2019 [26].

Comment: Education exposure was randomized by cluster (to the CA and their network members), but the role of each person not randomized CA/NM. Does that impact the analysis for any of the aims? Do the authors compare outcomes between CA and NM? If this does not matter, could the authors please explain this in the text?

Response: No, people were allowed to self-select into CA vs NM status. This does limit the generalizability of these results, but it does not impact the internal validity of this study. We do not directly compare outcomes between CAs and NMs, as all benefits of randomization would be lost here. We have added language to the paper explaining this.

Relevant Text: Here, we define CAs based on the potential for PLH to become so - by self-selecting into the study, PLH identify themselves as potential CAs. The NAMWEZA intervention is then designed to foster a CA's ability to truly act as a Change Agent, rather than in name only.

Comment: The authors describe the loss-to-follow-up. Was there any other information missing at baseline or follow-up (exposure, outcome, covariates)? If so, please describe. Perhaps the sentence about loss-to-follow-up could be moved closer to the discussion of the MCAR assumption.

Response: Thank you for this comment and recommendation. We have moved the sentence in question, and we believe it does flow better now. We have also added a sentence indicating that given a respondent was not LTFU, we have complete data on them.

Relevant Text: In this way, all demographic and contextual variables were measured at baseline. During this study, there was little loss-to-follow-up among the CAs (< 10%), but much higher loss among the NMs (36.8%) [31]. Given an NM or CA was not lost to follow-up, complete information was available on all additional variables, including exposure, outcome, and covariates.

Comment: For the permutation test, what are the assumptions to perform this test? Are they met in the analysis? Please clarify for the reader. Can the authors be more specific about the statistics they computed in the 1000 permuted samples?

Response: We have altered the phrasing to indicate exactly what we calculate in each sample. As a non-parametric test, no assumptions about the distributions are necessary as the distribution is drawn directly from the data. We have also indicated this in the paper.

Relevant Text: Because CAs and NMs self-selected into their respective group, and randomization only occurred within groups, differences between the groups were to be expected. However, we only examined homophily of baseline characteristics rather than of outcome, because comparison of outcomes between CAs and NMs would remove the benefits of randomization. We therefore assessed statistical significance of homophily on the set of CA-NM dyads using a permutation test, a non-parametric test which builds a distribution directly from the data. We then randomly permuted CA-NM ties (keeping number of ties per CA constant), and then recalculated the difference or percent concordant, respectively, 1,000 times.

Comment: Did the authors use one NM per CA for the mediation analysis? Or multiple NM per CAs? The sentence at the bottom of page 11 sounds like multiple NMs were used in the analysis. Does this violate any assumptions of that approach?

Response: Thank you for this question. Yes, we used multiple NMs per CA in the main analysis. Because this does violate the assumption of independence between observations, we also performed a sensitivity analysis, outlined on page 9. This removed 48NMs, leaving the point estimates nearly identical and also leaving significance levels unchanged. Confidence intervals were somewhat wider, however, due to the reduced sample size.

Relevant Text: One assumption of this analysis is that the dyads are independent, which is violated here; if a CA recruited more than one NM, the multiple CA-NM dyads involving the same CA would not be independent. To address this, we performed the analysis after randomly removing NMs until each CA had only a single NM. This resulted in removing 48 NMs, just 6.7% of the population. We found that the point estimates were nearly identical, but that the confidence intervals were slightly larger due to the reduced sample size. No coefficients changed from significant to non-significant in this analysis (data not shown).

Minor Comments

1. Introduction:

    a. Page 7, line 13: Do the authors mean to increase the reach of the intervention, rather than enrollment?
    b. Response: No, we meant enrolment here. By understanding the people who *didn't* join the intervention (among NMs who were PLH), and how they differed from CAs, we might better be able to identify ways to increase enrolment in the future.

2. Methods

    a. Page 7, line 31: What do the authors mean by "foster through the intervention?
    b. **Response:** We meant that we would CAs seeing themselves as true change agents as they progress through the intervention. We have clarified this in the text.

c. Page 8, line 27: What do the authors mean by "net of temporal or geographical trends"? Please clarify. Perhaps this is a point for the statistical methods section that relates to the methods used?

d. **Response:** Thank you for this comment. This was an unnecessarily-obtuse phrasing to indicate randomization. We have removed this sentence, as the intent appears clear even without it.

c. Were all demographic and contextual variables measured at baseline and prior to the exposure? Please add that point. For the effects estimated, it is important to have the intervention prior to the outcome and the covariates measured at baseline (prior to the intervention).

d. **Response:** Yes, they were all measured at baseline, and we have indicated this. The intervention was also administered prior to the outcome, but was not measured, *per se.* Rather, we knew whether the intervention had been administered via randomization tables of the participants.

e. Page 10, line 29: I would suggest saving the discussion of the limitations of hazard ratios for the discussion section. The second limitation is that it provides only one estimate for the duration of the study.

f. **Response:** We have edited this accordingly.

e. Should the assumptions be reported before the methods? What other assumptions are made in the three parts of the analysis? The second aim treats the network members as independent. Is the third aim assuming that network members and CA are exchangeable, perhaps conditional on covariates?

f. Response: This is a good point, and we have moved the assumption paragraphs after the description of the method. The third aim does not need to assume network members and CAs to be exchangeable (which they likely are not), but that intervened CAs and unintervened CAs are exchangeable, and likewise for NMS (both of these are expected but not guaranteed given randomization).

g. Page 10, line 40: The authors should clarify that these models are about loss to follow-up for the NM.

h. **Response:** We have done so.

3. Results

a. Does this result even need a statistical test: "Only 12.3% of NMs were HIV-positive, compared to all CAs"? How did the authors perform a statistical test if one of the cells was zero?

b. **Response:** This is a fair point, and we have removed the p-value from the table. Since this was a non-parametric test, in 0 of the 1,000 permutations was the group value as extreme as the observed value, hence the p-value. Had we performed a parametric test here, the CAs would have had a variance of 0, leading to an infinite t-score, leading to a p-value of 0. The permutation test allowed us to avoid this difficulty.

4. Discussion

    a. Page 13, line 47: Please clarify for which analyses the loss to follow-up may have biased. This bias is particularly a concern if the dropout was different across intervention arms and associated with the outcome.

    b. **Response:** Thank you for this comment, as we have now clarified that only the final spillover analysis may be affected by LTFU.

    c. Where is the discussion about contamination and a possible violation of the partial interference assumption?

    d. **Response:** Thank you for this comment. We have highlighted the discussion of the partial interference assumption in the methods section. By connecting it to the introduction of the assumption itself, we believe readers will not need to jump back and forth through the manuscript for this as they might have done if it were included in the discussion section.

5. Tables and Figures

    a. In the figure on page 23, what are the different arrows (dash vs. solid)?

    b. **Response:** The solid lines are standard arrows in a DAG. The dashed lines indicate which line or lines represent the effect in question. So the NDE is represented by the single causal arrow from A to Y_{NM}, but the NIE is represented by the union of the A to Y_{CA} and Y_{CA} to Y_{NM} lines. We have added language to the caption to make this clearer.

VERSION 4 – REVIEW

REVIEWER	Ashley Buchanan University of Rhode Island, US
REVIEW RETURNED	29-Jun-2020

GENERAL COMMENTS	General Comments
	The manuscript is much improved and many of these technical points are now clear. The paper reads nicely and will be an important contribution to the literature. I thank the authors for engaging with these comments and improving their paper. The authors clarified their methods approach and possible limitations, provided interesting possible explanations of the results, and clarified aspects of the study design. I included one major comment and some additional minor comments below.
	The authors now clearly explain their approach using mediation and why this is valid, including the sensitivity analysis with only one NM per CA, which seems critical to their approach. I am not sure that page 402-403 of Vanderweele's book is advocating for using mediation to evaluate spillover; however, the book by Hong certainly does in Chapter 15. I would suggest that the authors cite the book by Hong (2015). The authors may want to review this chapter and ensure their approach aligns with this framework, including the assumptions on page 396. In this analysis, was the mediator (CA knowledge change) randomized? If not, what adjustments were made for possible confounding of the mediator? I do not yet see this reported in the paper. The authors should make very clear up front

that this mediation framework for spillover only applies to dyads when the identification assumptions are met.

On page 8, do the authors mean "randomization at the group level", rather than "within groups". Please clarify. My understanding is that the CA-NM clusters were randomized to the intervention, but the status of CA/NM was not randomized, as the CA came forward themselves and CA knowledge was not randomized (the mediator). On page 8, the authors say "similarly randomized" and this could also be clarified.

For the point about the permutation test with zero cell counts, if the authors ensure the test is valid, then this is fine to report in the paper. The nonparametric test may be better described as "has no distributional assumptions", rather than "builds a distribution directly from the data".

I found the explanation of CA to "act as CA, rather than name only" a little confusing. Could the authors please clarify this point?

The authors may want to carefully check the literature and possibly cite Benjamin-Chung (2017) before claiming that spillover of HIV knowledge is yet unknown in SSA.

On page 5, instead of "given a survey", "completed a survey" seems more accurate. On page 8, rather than "passing the intervention", "sharing the intervention" sounds better to my ear.

When referring to Chronbach's alpha, the authors likely mean the "questions on the measure" not "the measure".

For "Complete HIV Knowledge", please define the other group, which I believe is "got at least one incorrect".

Throughout, please review carefully for tense. Any analyses or study already conducted should be in past tense.

For the statement about not sharing data, the authors may want to give a stronger reason.
There are many studies about HIV that now share their data (e.g., ACTG).

In the text, the first time RR is mentioned, please write out as "Risk Ratio".

References:

Benjamin-Chung, J., Abedin, J., Berger, D., Clark, A., Jimenez, V., Konagaya, E., ... & Miguel, E. (2017). Spillover effects on health outcomes in low-and middle-income countries: a systematic review. International Journal of Epidemiology, 46(4), 1251-1276.

Hong, G. (2015). Causality in a social world: Moderation, mediation and spill-over. John Wiley & Sons.

VERSION 4 – AUTHOR RESPONSE

Major comments:

Comment: I am not sure that page 402-403 of Vanderweele's book is advocating for using mediation to evaluate spillover; however, the book by Hong certainly does in Chapter 15. I would suggest that the authors cite the book by Hong (2015). The authors may want to review this chapter and ensure their approach aligns with this framework, including the assumptions on page 396. In this analysis, was the mediator (CA knowledge change) randomized? If not, what adjustments were made for possible confounding of the mediator? I do not yet see this reported in the paper. The authors should make very clear up front that this mediation framework for spillover only applies to dyads when the identification assumptions are met.

Response: We thank you for making these points. We however don't think that the book by Hong more clearly advocates for spillover-by-mediation-analysis than the vanderWeele book does. For example, the latter states the following: "As we go through this chapter, we will see that many of the methods and approaches for mediation and interaction have direct analogues in the spillover effect context" (pp397, vanderWeele). Moreover, the examples in vanderWeele's book more clearly align with the context of our study. In the driving example in Hong's book, *all people* are exposed to the intervention in an exposed cluster. In vanderWeele's book, only some people in each cluster are exposed, and the spillover is the subsequent effect on the unexposed in each cluster. As each CA-NM pair forms a cluster in this study, with the CA exposed and the NM not directly exposed to the intervention, the setting in vanderWeele's book is more apt, and we believe that interested readers will be more easily able to translate the context of our study to the formulas and description laid out by vanderWeele.

All that said, we do believe the book by Hong is a valuable resource, and we have added it as a reference. We also recognize that regardless of the primary source, the assumption of no confounding between the outcome of the CA and the outcome of the NM is not met if the pairs are not randomized (assumption 4 in section 15.2.1 in Hong, assumptions A15.3 and A15.6 on page 414 in vanderWeele). In our study, these pairs are not randomized, so the assumption is not expected to hold here. We do adjust for a wide variety of potential confounders (largely the variables used in the loss-to-follow-up analysis) to account for this possibility. We have added some detail tthe text to indicate this.

Relevant Text:

Importantly, this analysis requires a number of assumptions and applies to dyads only when these assumptions are met.
A third assumption of this analysis is that the outcomes of the CA and NM are independent conditional on the CA's exposure, or conditional on the CA's exposure and other confounding variables [42,46]. Because CA-NM pairs self-select and are not randomized, we do not expect these outcomes to be independent conditional of the CA's exposure, and so we adjust for additional variables to meet this assumption. In our analysis, we therefore adjust for all variables used in the log-binomial regression.

[46] Hong G, 2015. *Causality in a social world: Moderation, mediation and spill-over.* John Wiley & Sons.

Minor Comments

Comment: On page 8, do the authors mean "randomization at the group level", rather than "within groups". Please clarify. My understanding is that the CA-NM clusters were randomized to the intervention, but the status of CA/NM was not randomized, as the CA came forward themselves and CA knowledge was not randomized (the mediator). On page 8, the authors say "similarly randomized" and this could also be clarified.

Response: We have clarified both instances mentioned to make it explicitly clear that we were primarily randomizing when the CA received the intervention, and treat the NM as being exposed to the intervention along with their respective CA.

Relevant Text:

Because CAs and NMs self-selected into their respective group, and only CAs were directly randomized (with their NMs being randomized along with them), differences between the groups were to be expected.
As the wedge in which the CA received the *NAMWEZA* intervention was randomized, we treat each NM as being randomized to exposure to *NAMWEZA* at the same time as their CA.

Comment: For the point about the permutation test with zero cell counts, if the authors ensure the test is valid, then this is fine to report in the paper. The nonparametric test may be better described as "has no distributional assumptions", rather than "builds a distribution directly from the data".

Response: We thank the reviewer for this alternative phrasing – we concur that this is a better description and have changed it accordingly.

Relevant Text:
We therefore assessed statistical significance of homophily on the set of CA-NM dyads using a permutation test, a non-parametric test which has no distributional assumptions.

Comment: I found the explanation of CA to "act as CA, rather than name only" a little confusing. Could the authors please clarify this point?

Response: We appreciate this comment, and have rewritten it to make it more clear.

Relevant Text:
Although we refer to them as 'CAs' throughout, participants in the trial enrolled with varying levels of ability to act as a Change Agent. Through receiving the NAMWEZA intervention, we hypothesize that CAs will be able to truly self-actualize and subsequently act as Change Agents in their community. Although the intervention does not increase their HIV knowledge, it is still useful to the CAs, as it empowers them to act as CAs in their community.

Comment: The authors may want to carefully check the literature and possibly cite Benjamin-Chung (2017) before claiming that spillover of HIV knowledge is yet unknown in SSA.

Response: We thank the reviewer for asserting that we check this assertion. On examination, we do generally stand by our statement, as neither the mentioned review article or articles published subsequently directly examine spillover of HIV knowledge in a controlled setting that make the spillover identifiable. However, we did find research examining spillover through proxies (e.g. inviting social network members to watch the same program or time spent shopping a market). Because there has been some examination of spillover (albeit not as directly as measured in the present study), we have toned down the relevant language to reflect this, and cited additional references.

Relevant Text:
Studies have also used proxy variables for social network ties such as inviting social network members to watch educational programming [13] or time spent shopping at the market [14] to evaluate spillover effects for HIV knowledge, generally finding evidence for spillover. However, spillover in HIV knowledge between known social network ties generally remains understudied, particularly in sub-Saharan Africa [15]. We therefore aim to determine whether social network members those receiving an HIV behavioral/knowledge intervention also increase their HIV knowledge.

[13] Banerjee A, La Ferrara E, and Orozco-Olvera VH, 2019. The entertaining way to behavioral

change: Fighting HIV with MTV. The World Bank.

[14] Self S, and Grabowski R, 2018. Factors influencing knowledge of HIV/AIDS in Nepal: role of socioeconomic interactions. Journal of Social and Economic Development, 20(1), pp.17-191.

[15] Benjamin-Chung J, Abedin J, Berger D, Clark A, Jimenez V, Konagaya E, Tran D, Arnold BF, Hubbard AE, Luby SP, and Miguel E, 2017. Spillover effects on health outcomes in low-and middle-income countries: a systematic review. International journal of epidemiology, 46(4), pp.1251-1276.

Comment: On page 5, instead of "given a survey", "completed a survey" seems more accurate. On page 8, rather than "passing the intervention", "sharing the intervention" sounds better to my ear.

Response: We have changed these accordingly.

Relevant Text:
Participants completed a baseline questionnaire and were randomized to one of three waves in which to receive the intervention.
Finally, since the trial showed beneficial effects on the HIV knowledge of the NMs following the intervention [27], we aimed to elucidate exactly what caused the HIV knowledge of NMs to increase - either CAs gaining knowledge through the intervention and sharing it, or the CAs being empowered by the intervention to share existing knowledge.

Comment: When referring to Chronbach's alpha, the authors likely mean the "questions on the measure" not "the measure".

Response: Yes, this is a more-accurate way to state this, and we have changed it accordingly.

Relevant Text:
In these populations, questions on the measure had a Chronbach's alpha of 0.78.

Comment: Throughout, please review carefully for tense. Any analyses or study already conducted should be in past tense.

Response: We have gone through and edited this accordingly.

Relevant Text: Throughout

Comment: For the statement about not sharing data, the authors may want to give a stronger reason. There are many studies about HIV that now share their data (e.g., ACTG).

Response: This is a valuable point. We have added additional detail indicating that the networked nature of the data adds an additional vulnerability were the data to be

Relevant Text:
The data are not publicly available due to the sensitive nature of HIV infection status and the socially-networked nature of the data. Because the data include specific information on social ties, some of whom have not disclosed HIV sero-status to one another, the risks of individual identification and compromising HIV sero-status are greatly increased.

Comment: In the text, the first time RR is mentioned, please write out as "Risk Ratio".

Response: We have done so.

Relevant Text:
Risk ratios (RRs) obtained via log-
binomial regression indicated that characteristics of both the CAs and the NMs significantly predicted l
oss to follow-up (Table 2).

VERSION 5 – REVIEW

REVIEWER	Ashley Buchanan University of Rhode Island, USA
REVIEW RETURNED	03-Aug-2020

GENERAL COMMENTS	The authors have sufficiently addressed the comments. I included some minor points of clarification below. This is an interesting paper and a great contribution to the literature! On page 9 when describing the mediation analysis, please explain what was adjusted for in each model (model of randomized intervention on mediator and model of mediator on outcome). The authors describe this when discussing the third assumption, but it would be best to clarify that this was the adjustment made for the mediation analysis. The authors could clarify the time ordering between the randomized intervention, the mediator and the outcome. I believe the randomized intervention happens at each wave, then the CA randomized status determines if the CA-NM was intervention or control (by wave 1), then the follow-up visit for the CA (the mediator) occurs by the end of wave 1. Also, by the end of wave 1, the follow-up visit for the NM occurs. This could be written more clearly in the methods section. Does the mediator come before the NM outcome in time, or is there possibly some overlap there? If there is not a clear ordering of exposure, mediator and outcome in some cases, perhaps list as a limitation? On a minor note, what does "directly randomized" mean, perhaps use a different wording here, such as "only CA were randomized"?

VERSION 5 – AUTHOR RESPONSE

Response to Reviewers

Comment: On page 9 when describing the mediation analysis, please explain what was adjusted for in each model (model of randomized intervention on mediator and model of mediator on outcome). The authors describe this when discussing the third assumption, but it would be best to clarify that this was the adjustment made for the mediation analysis.

Response: Thank you for asking for this clarification – it allows us to explain exactly what we controlled for, and to point to the fact that randomization is not a cure-all, as selection bias is always a possibility.

Relevant Text: In the models estimating the effect of the exposure on the mediator and estimating the effect of the mediator on the outcome, we adjusted for all the variables included in our logistic regression above. Although the randomization of the exposure minimized some potential bias, the loss-to-follow-up among the CAs indicates that selection bias could remain a concern, so we control for the variables which may also impact loss-to-follow-up.

Comment: The authors could clarify the time ordering between the randomized intervention, the mediator and the outcome. I believe the randomized intervention happens at each wave, then the CA randomized status determines if the CA-NM was intervention or control (by wave 1), then the follow-up visit for the CA (the mediator) occurs by the end of wave 1. Also, by the end of wave 1, the follow-up visit for the NM occurs. This could be written more clearly in the methods section. Does the mediator come before the NM outcome in time, or is there possibly some overlap there? If there is not a clear ordering of exposure, mediator and outcome in some cases, perhaps list as a limitation?

Response: This was a useful statement to make and explicitly delineate our randomization definition, as well as the timing of interviews.

Relevant Text: In other words, the CAs randomized to receive NAMWEZA during the first wave would have potentially indirectly exposed their NMs to the intervention when the NM completed their follow-up questionnaire after Wave 1. Therefore, the NMs were divided into "exposed" (N=381) and "unexposed" (N=329) groups based on whether their respective CA was randomized into receiving NAMWEZA during the first wave or not. The CAs always completed their Wave 1 follow-up interview before their NMs were invited to complete their Wave 1 follow-up interview.

Comment: On a minor note, what does "directly randomized" mean, perhaps use a different wording here, such as "only CA were randomized"?

Response: We meant this to disambiguate the CAs who are randomized, vs. NMs who were just "along for the ride", but understand that this did not actually clarify the point. We have done as suggested.

Relevant Text: Because CAs and NMs self-selected into their respective group, and only CAs were randomized (with their NMs being randomized along with them), differences between the groups were to be expected.