

# BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email [info.bmjopen@bmj.com](mailto:info.bmjopen@bmj.com)

# BMJ Open

## Network Spillover Effects and Follow-Up Correlates in a HIV Prevention Intervention in Tanzania

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-033759
Article Type:	Research
Date Submitted by the Author:	21-Aug-2019
Complete List of Authors:	Lienert, Jeffrey; University of Pennsylvania Perelman School of Medicine, Center for Health Care Innovation; Philadelphia VAMC, CHERP Fawzi, Mary C; Harvard Medical School, Global Health and Social Medicine McAdam, Keith; London School of Hygiene and Tropical Medicine, Epidemiology Kaaya, Sylvia; Muhimbili University of Health and Allied Sciences, School of Medicine Liu, Yuanyuan Todd, Jim; London School of Hygiene and Tropical Medicine, Population Health Andrew, Irene; Management and Development for Health, Onnela, Jukka Pekka; Harvard T.H. Chan School of Public Health, Biostatistics
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, STATISTICS & RESEARCH METHODS

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

## Network Spillover Effects and Follow-Up Correlates in a HIV Prevention Intervention in Tanzania

Jeffrey Lienert<sup>a,b,j</sup>, Mary C. Smith Fawzi<sup>c</sup>, Keith McAdam<sup>d</sup>, Sylvia Kaaya<sup>e</sup>, Yuanyuan Liu<sup>i</sup>, Jim Todd<sup>g</sup>, Irene Andrew Lema<sup>f</sup>, Jukka-Pekka Onnela<sup>h</sup>

<sup>a</sup>Perelman School of Medicine, University of Pennsylvania <sup>b</sup>Center for Health Equity Research and Promotion, Philadelphia VA <sup>c</sup>Department of Global Health and Social Medicine, Harvard Medical School <sup>d</sup>Department of Clinical Tropical Medicine, London School of Hygiene and Tropical Medicine <sup>e</sup>Department of Psychiatry and Mental Health, Muhimbili University of Health and Allied Sciences <sup>f</sup>Management and Development for Health <sup>g</sup>Department of Population Health, London School of Hygiene and Tropical Medicine <sup>h</sup>Department of Biostatistics, Harvard T. H. Chan School of Public Health <sup>i</sup>University of Texas School of Public Health <sup>j</sup>Corresponding Author: Jeffrey Lienert, [jeffrey.lienert@penntmedicine.upenn.edu](mailto:jeffrey.lienert@penntmedicine.upenn.edu)

## Abstract

**Objectives:** We aim to describe the social network members of participants of a behavioral intervention, and examine how the effects of the intervention may spill over into these network members.

**Design:** Secondary analysis of a step-wedge randomized controlled trial.

**Setting:** Change Agents (CAs) were recruited from waiting rooms of HIV treatment facilities in Dar es Salaam, Tanzania, and their Network Members (NMs) were recruited directly by CAs.

**Participants:** We enrolled 662 CAs in an HIV behavioral intervention. They, along with 710 of their NMs, completed baseline and follow-up interviews from 2011 to 2013.

**Primary and Secondary Outcomes:** The primary outcome of this study was change in NMs' HIV knowledge, and the secondary outcome was whether the NM was lost to follow-up.

**Results:** At baseline, many characteristics were different between NMs and CAs. We found a number of NM characteristics significantly associated with follow-up of NMs, particularly female gender (OR=1.64) and HIV knowledge (OR=20.0); only one CA variable was significantly associated with NM follow-up: having a private source of water (OR=2.17). The 14.2% increase in NMs' HIV knowledge was largely due merely to CAs participating in the intervention, rather than CAs transmitting *new* knowledge to their NMs.

**Conclusions:** Characteristics of social network members of PLH may play a role in study retention. Additionally, the HIV knowledge of these NMs increased largely as a function of CA participation in the intervention, indicating that intervening in highly-connected individuals may maximize benefits to the potential population for whom spillover can occur.

*Keywords:*

mediation, natural direct effect, natural indirect effect, follow-up, retention, hidden population

## 1. Strengths and Limitations of This Study

- The recruitment method, focused on social network members of persons living with HIV, allowed us to access an at-risk population not easily accessible via other means.
- The study design of the trial in which this study is nested was an ideal situation in which to employ a mediation analysis that informs our understanding of how this and similar trials may spillover in a population.
- Although the high dropout rate of social network members meant we could examine factors leading to dropout, it exposed the outcomes to bias if those who dropped out were different from those who did not.

## 2. Introduction

Despite recent decreases in mortality, HIV/AIDS is currently the ninth leading cause of years of life lost (YLL) globally [1]. Research into eliminating HIV has focused on two fronts: biological and behavioral, which combined have decreased the YLL due to HIV by an estimated 51% [1]. Behavioral interventions seek to curb transmission of HIV through reduction of risk behavior including safer sexual practice and injection drug use [2, 3]. Behavioral change also impacts the effectiveness of the biological methods, given that low adherence will compromise their effectiveness [4]. However, neither method is fully effective without the knowledge of these treatments and preventative behaviors. It is therefore important to increase HIV-related knowledge as a necessary, but not sufficient, step towards achieving the 90-90-90 treatment goals and ultimately eliminate HIV [5].

The importance of adequate HIV knowledge has been recognized for a period of time, as it is often a prerequisite for behavioral change [6]. It should be noted that increased knowledge does not necessarily translate into behavior change, as other factors such as motivation and skills also play a role [6]. Despite this, many trials have been conducted in persons living with HIV (PLH) which have increased HIV knowledge as an endpoint and have often found positive effects of interventions on HIV knowledge [7, 8].

Knowledge gained by participants in these trials can also be freely shared with members of their social network, in what is known as a spillover effect [9]. Specifically, a spillover effect is a change to an individ-

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

ual's behavior that is due to an exogenous change to another's. In the context of an intervention, it means a change to a non-participant in the intervention stemming from someone else's participation in the intervention. What remains unknown, however, is whether or not a spillover effect exists for HIV knowledge during and after an intervention.

New knowledge can come from a variety of sources, one of the most important of which is a person's social network [10, 11]. Social networks are of particular import because new knowledge can lead to cascades of behavior change, where people subsequently educate those in their social network, in what is known as social influence [12, 13, 14]. This has been directly examined in participant-driven interventions, where early participants directly educate members of their social network [15]. Characteristics such as knowing the HIV status of network members has been shown to be the most important predictor of engaging in prevention advocacy [16]. Work on diffusion through social networks has shown that spreading intervention effects beyond the initial population increases the cost-effectiveness of these interventions [17]. These findings imply that certain aspects of knowledge or behavior may spread more or less efficiently through networks comprising individuals with specific characteristics, which may need to be accounted for in network interventions.

Additionally, the networks of PLH are often hidden networks [18], due to the continued stigma of HIV and AIDS in many settings [19]. Because of this, if a PLH or person at-risk of acquiring HIV does not want to participate in an intervention, there is little recourse other than information transmitted via social networks, or targeted sampling techniques which are not always effective [20]. This is particularly important in low- and middle-income countries, as a recent systematic review found only 54 studies researching spillover effects [21]. Therefore, understanding exactly how information spreads from participants in an intervention to members of their social network, who may be largely inaccessible via other means, is important for reaching the greatest number of people about HIV prevention. Understanding what makes these persons different from those who enroll in the intervention itself is important, as it may point to ways in which to increase enrollment of these populations.

Based on the above gaps in the literature, we conducted a study on network members of PLH enrolled in the larger *NAMWEZA* intervention [22]. The trial recruited PLH to serve as Change Agents (CAs) and to reach out to their social network members (NMs) about knowledge of HIV and safer sexual practices [23].

Our goals in this study were threefold: 1) to understand how the NMs differed from their CAs, 2) understand correlates of dropout for the NMs, and 3) to understand by what method HIV knowledge transferred to the NMs from the CAs. Understanding how the information and behaviors transfer will allow HIV researchers and others to take advantage of this knowledge and improve upon prevention interventions in the future.

### 3. Methods

#### 3.1. Study Population

We analyzed social network data from the Agents of Change trial [24], which was a stepped-wedge randomized controlled trial that enrolled PLH to become Change Agents (CAs) by informing members of their social network (NMs) about knowledge of HIV and safer sexual practices [25]. CAs were recruited from the waiting rooms of HIV care and treatment centers in Dar es Salaam, Tanzania. Participants were given a baseline questionnaire and were randomized to one of three waves in which to receive the intervention. The intervention comprised 10 weekly structured sessions aimed at empowering PLH to become HIV prevention change agents in their communities. The sessions aimed to increase psychosocial and communication skills through an Appreciative Inquiry approach [26]. Within one month of each wave of the intervention, CAs and NMs were given follow-up surveys. The study lasted from November 2010 to January 2014.

At baseline, participants were also asked to recruit members of their social networks who they felt were at particularly high risk of contracting or spreading HIV. These network members could be either HIV positive or negative. Upon successful recruitment, these network members were also given a baseline survey. The NM was only aware the CA was a participant in the intervention if the CA shared this information with them, which many did not due to HIV-related stigma [27]. Each CA therefore formed a CAN-NM pair with each NM they recruited. Inclusion criteria for the follow-up analyses were for CAs who completed at least one follow-up interview and had at least one NM who completed a follow-up questionnaire. In this way, outcomes could be computed for the CA and NM of each dyad.

#### 3.2. HIV knowledge

To assess HIV knowledge of CAs and NMs, we used the brief HIV knowledge questionnaire [28]. This scale comprises 18 questions, which focus on an individual's knowledge of how HIV can spread and other



1  
2  
3  
4  
5 characteristics of the virus and AIDS. This knowledge is crucial for an individual's subsequent safe-sex  
6 practices to reduce the risk of transmission. This instrument has been used previously in sub-Saharan Africa  
7 and has demonstrated reasonable reliability of 0.75 Cronbach's alpha in South Africa [29], whereas it was  
8 0.78 in the original study. It has also been translated to Swahili, with only minor differential item functioning  
9 [30]. This indicates that the measure performs adequately in other, similar populations, relevant to this  
10 work. Because the majority of CAs and NMs answered all questions correctly, we summarize this measure  
11 as "Complete HIV Knowledge", a dichotomous variable for whether the participant correctly answered all  
12 questions.  
13  
14  
15  
16  
17  
18

### 19 3.3. Statistical Analysis

20  
21 The first analysis was to assess homophily, defined as the extent to which CAs and NMs were similar to  
22 one another in terms of a number of sociodemographic characteristics and HIV-related risk factors. Because  
23 CAs could have more than one NM, we assessed significance using univariate intercept-only Generalized  
24 Estimating Equations (GEE) with a Normal or Bernoulli distribution for continuous and dichotomous vari-  
25 ables, respectively. We used exchangeable working correlation structures in these models. Analyses were  
26 run using R v3.1.1.  
27  
28  
29  
30

31 To accomplish our second aim of understanding what was associated with NMs completing all their  
32 follow-up interviews, we fit a logistic regression to determine predictors of follow-up. In order to examine  
33 the association between these same variables and the time to follow-up, we also employed a Cox regression.  
34 In the Cox regression, start of follow-up was defined as the time at which the NM completed their baseline  
35 interview. The outcome was loss-to-follow-up. NMs were interval-censored between the date of their last  
36 completed interview and last available date of the next scheduled interview, as NMs could have decided not  
37 to continue to participate at any point during that time.  
38  
39  
40  
41  
42

43 Finally, since the trial showed beneficial effects on the HIV knowledge of the NMs following the inter-  
44 vention [23], we aimed to elucidate the putative causal mechanism through which the CAs impacted their  
45 network members. As shown by (author?) [31], social network spillover effects can be broken down into  
46 direct and indirect effects in the case of dyadic relationships (Figure 1). This method has since been used  
47 for novel spillover analyses [32]. Although previous studies showed that this same type of analysis cannot  
48  
49  
50  
51  
52

1  
2  
3  
4  
5 be done on full network data, the data in this study consisted of only paired individuals, the CAs and their  
6 NMs, so in this case the analysis does not result in biased estimates [33, 34].

7  
8 The method parses social influence into direct and indirect effects. The natural indirect effect (NIE) is  
9 the effect by which the intervention changes the CA's HIV knowledge, which then changes the NM's HIV  
10 knowledge. The NIE is similar to the effect observed in many participant-driven interventions: an initial  
11 participant receives the intervention, increasing their knowledge, and they subsequently pass their increased  
12 knowledge to members of their social network [15]. The natural direct effect (NDE) is the effect merely re-  
13 ceiving an intervention has on an NMs outcome, irrespective of the CA's outcome (in this case HIV knowl-  
14 edge). For instance, this could occur if all CAs begin with good knowledge of HIV, and the intervention  
15 empowers them to convey knowledge they already had to their NMs. In order to estimate these effects, the  
16 published SAS macro developed by VanderWeele and colleagues was used, and analyses were run using  
17 SAS v9.2 (SAS Institute, Cary, NC) [31]. The data are not publicly available due to the sensitive nature of  
18 HIV infection status.  
19  
20  
21  
22  
23  
24  
25  
26  
27

### 28 *3.4. Patient and Public Involvement*

29 Patients were not involved in the design of this study. As part of the recruitment process, patients were  
30 instructed to recruit members of their social networks who they felt were at high risk of HIV infection - in  
31 this way participants could guide the recruitment of NMs to those most at-risk. We currently have no plans  
32 to disseminate the results of this study to participants. Participants assessed the burden of the intervention  
33 via qualitative interviews; we found that many felt the timing was burdensome.  
34  
35  
36  
37  
38

## 39 **4. Results**

40  
41 The *NAMWEZA* study recruited 662 Change agents (CA), of whom 453 (68.4%) completed at least  
42 one follow-up questionnaire. These 662 CAs in turn recruited 710 network members (NM) who took the  
43 baseline questionnaire. Of the 710 NMs, 449 (63.2%) also completed a follow-up questionnaire (Table 1).  
44 At baseline, CAs were on average older than their NMs. More of the NMs were employed than their CAs  
45 (69.3% vs. 55.0%,  $p < 0.0001$ ), but were less likely to have at least 7 years of education (52.0% vs. 52.3%,  
46  
47  
48  
49  
50  
51  
52

p<0.0001). Only 12.3% of NMs were HIV-positive, compared to *all* CAs (p<0.0001). Complete data was obtained at baseline for all CAs and NMs.

Characteristic	Number of NMs (%) or Mean (SD) (N=710)	Number of CAs (%) or Mean (SD) (N=662)	P-value
Age	33.0 (11.1)	38.9 (9.7)	<0.0001
Female	380 (53.7%)	349 (53.9%)	0.956
Employed	490 (69.3%)	356 (55.0%)	<0.0001
At least 7 years education	369 (52.0%)	584 (82.3%)	<0.0001
Complete HIV knowledge	638 (89.9%)	598 (90.4%)	0.78
Persons sleeping in home	5.00 (5.77)	3.87 (3.72)	<0.0001
Married	373 (52.7%)	338 (51.1%)	0.61
HIV Positive	87 (12.3%)	662 (100%)	<0.0001
Private source of water	309 (43.7%)	263 (39.7%)	0.20

Table 1: Demographic characteristics at baseline, with the results of an intercept-only GEE for differences.

Logistic regression showed that characteristics of both the CAs and the NMs significantly predicted loss to follow-up (Table 2). The NM being female, having complete HIV knowledge, being employed, and being married were all significant predictors of increased odds of being followed-up. Each additional person sleeping in the home of the NM per room used for sleeping reduced the odds of follow-up (OR=0.81, 95% CI: 0.67,0.98) as did the NM living with HIV (OR=0.42, 95% CI: 0.18,0.99). CA having a private water source was significantly associated with increased odds of the NM being followed-up (OR=2.17, 95% CI: 1.33,3.57), even after controlling for the NM having a private source of water.

The Cox proportional hazard model showed similar results to the logistic regression, but with fewer significant results. Only the NM being married and the NM living with HIV significantly predicted the time-to-follow-up of the NMs. Additionally, all of the hazard ratios were closer to the null relative to the corresponding odds ratios.

We assess the mechanism of spillover effects on HIV knowledge, the main effect of the intervention on the NMs. The analysis showed that NMs who were exposed to treatment via their CA had an increase in HIV knowledge of 12.9% on average (95% CI: 0.06, 0.20) from a baseline percentage of 78% [23]. This was broken down into a natural indirect effect (NIE) of 0.6% (95% CI: -0.0060,0.02), which is the effect the intervention had via the CA's HIV knowledge changing, and a natural direct effect (NDE) of 12.3% (95%

Characteristic (N=459)	Adjusted OR (95% CI)	Adjusted HR (95% CI)
Characteristics of NMs		
Gender (Female)	1.64* (1.02,2.63)	1.18 (0.94,1.50)
Difference in age of NM and CA (per year)	1.01 (0.99,1.03)	1.01 (0.995,1.02)
Age of NM	0.98 (0.95,1.02)	0.99 (0.97,1.04)
Complete HIV knowledge (vs. none)	20* (3.70,125)	2.20 (0.97,5.01)
Employed	1.5* (1.04,2.5)	1.15 (0.89,1.50)
Each additional person sleeping in home per room used for sleeping	0.81* (0.67,0.98)	0.92 (0.83,1.01)
Married	1.72* (1.04,2.86)	1.28* (1.01,1.64)
Living with HIV	0.42* (0.18,0.99)	0.71* (0.51,0.99)
Having a private source of water	0.86 (0.54,1.37)	0.97 (0.77,1.22)
Characteristics of CAs		
Gender (Female)	1.27 (0.76,2.08)	1.07 (0.84,1.37)
Complete HIV knowledge (vs. none)	0.36 (0.07,2.04)	0.64 (0.29,1.43)
Having a private source of water	2.17* (1.33,3.57)	0.97 (0.77,1.22)
Being employed	1.54 (0.95,2.50)	1.14 (0.90,1.43)
Being married	1.25 (0.77,2.04)	1.11 (0.88,1.43)

Table 2: Results of multivariate logistic regression and Cox-proportional hazard models on the dichotomous outcomes of whether the NM was followed-up, and the continuous outcome of time-to-follow-up, respectively. “\*” indicates significant at the  $p < 0.05$  level.

CI: 0.06,0.19), which is the effect the CA’s participation had on the NM’s HIV knowledge, irrespective of the CA’s HIV knowledge.

## 5. Discussion

Based on results from the *NAMWEZA* trial in Tanzania, we have shown a number of novel findings regarding the network members of HIV intervention participants, correlates of their retention in the study, and the spillover of HIV knowledge from CAs to NMs. These findings can inform the design of interventions in the future to maximally enroll and retain participants and ensure that as much information as possible is transmitted from the study participants to members of their social networks.

We found many significant differences between CAs and the NMs they recruited. On average, CAs were older, less likely to be employed, and more educated. All of this suggests that the NMs were being recruited from populations that were substantially different for the population from which the CAs came [9]. This may in part be due to the fact that when recruiting NMs, CAs were instructed to recruit those who may be at particular risk of contracting or spreading HIV. Therefore, CAs likely did not pick random members of their

1  
2  
3  
4  
5 social network, but those who were at high risk. This finding corroborates the claim by Latkin et al. (2013)  
6 that sampling via enumeration of network members by initial participants recruits a more diverse sample  
7 [35]. Additionally, the lower likelihood of employment of the CAs indicates they may have had more time  
8 for participating in the intervention [36]. This means that future studies either need to focus recruitment on  
9 employed persons, or encourage spillover from study participants to network members.

10  
11  
12  
13 Although the study design potentially accessed a separate slice of the population than other methods  
14 vis-à-vis the population of those at risk for contracting HIV, it may have been at least partially responsible  
15 for the large loss of follow-up of the NMs. It is important to note that more NM characteristics than CA  
16 characteristics were significantly predictive of follow-up in both the logistic and Cox proportional hazards  
17 models. Had CAs and NMs been more similar, and any of these characteristics acted in a causal fashion,  
18 the CAs may have had greater influence over their NMs, resulting in greater follow-up of the NMs due to  
19 encouragement from their CAs. We also note that less HIV knowledge and living with HIV predicted greater  
20 loss to follow-up among NMs, which means that those who might have benefited most from spillover of the  
21 intervention were more likely to discontinue their involvement. This does not mean that they did not receive  
22 any spillover, only that it was not recorded. Our estimate of the magnitude of the spillover may therefore be  
23 biased towards the null. This is also problematic more generally for interventions of this nature as the very  
24 people the intervention aims to benefit may not stay with the program.

25  
26  
27  
28  
29  
30  
31  
32  
33 The HIV knowledge gain experienced by the NMs was largely due to the NDE - i.e. knowledge spilled-  
34 over as a result of the CAs participating in *NAMWEZA*, not because the CAs' HIV knowledge increased,  
35 and they passed this new knowledge to their NM(s). This is most likely because HIV knowledge of all  
36 CAs started off relatively high (80%), so increases in HIV knowledge based on our questionnaire was not  
37 possible for all CAs. Following the intervention, CAs' average HIV knowledge did not significantly increase.  
38 Therefore, what prompted the increased knowledge of the NMs was likely merely their association with the  
39 CA. So instead of the CA providing informational support by directly sharing novel information [37], CAs  
40 may have felt more empowered to share their existing knowledge as a result of the intervention.

41  
42  
43  
44  
45  
46  
47 This finding is important for future interventions: spillover effects of this intervention will likely carry  
48 over only one degree of separation from the CAs (opposed to spreading indefinitely in a snowball effect),  
49 as the most important thing in transmitting HIV knowledge is receiving the intervention, which only CAs  
50  
51

1  
2  
3  
4  
5 receive. Because NMs do not actually receive the intervention, but only the contact with their CA, it is  
6 unlikely that they would feel empowered to become CAs themselves, thus limiting the spread of the inter-  
7 vention. This may give insight into how to design interventions in the future; if one wants to maximize the  
8 number of people who benefit, choosing CAs who are highly connected in the community would maximize  
9 the potential number of links by which spillover can occur. Alternatively, interventions can be designed to  
10 be self-propagating; if CAs are empowered to deliver the intervention to others, the effects seen here could  
11 spillover continuously.

12  
13 Our study is not without limitations. First, there was considerable loss to follow-up of the NMs. Al-  
14 though this actually informed our analysis of the correlates of loss to follow-up, it meant that our final  
15 analysis may have biased our results. Even though the exposure was randomized, the loss to follow-up can  
16 result in selection bias if the NMs who left the study were systematically different from those who remained.  
17 As we show, the NMs who dropped out were those who would have benefited the most from the intervention  
18 because they were more vulnerable and at higher risk than those who did not drop out. Because they are  
19 likely to have greater benefits from any spillover effects, we would expect this non-random loss to follow-up  
20 to result in an underestimate of the impact of the intervention in the network members. Second, our data  
21 did not perfectly fit the requirements of the causal mediation analysis: they were not entirely independent,  
22 since multiple NM-CA dyads shared a CA. However, when we randomly removed dyads until there were  
23 no repeated CAs, the results were qualitatively very similar, indicating that lack of independence did not  
24 unduly affect our results. Third, although we were able to tease apart the direct and indirect effects, we are  
25 unable to determine the mechanism of the natural direct effect; the data do not allow us to specify whether  
26 NMs increased knowledge through speaking to their knowledgeable CA, through researching HIV on their  
27 own, or some other mechanism. Future work will have to be done to examine these different pathways.

## 42 43 **6. Conclusions**

44  
45 These results have implications for the potential scale-up of the *NAMWEZA* intervention, as well as  
46 future studies and interventions that focus on behavioral interventions in social networks. First, our findings  
47 of minimal similarity between CAs and their NMs indicate that this recruitment method allows us to enroll  
48 participants from portions of the population that are not represented by the CAs alone. Understanding  
49  
50  
51

1  
2  
3  
4  
5 how this then translates into closeness of ties between NMs and CAs and the implications for follow-up  
6 may inform strategies for retention and may increase the impact of future HIV behavioral interventions.  
7 The mediation analysis presents a compelling picture of how best to ensure the benefits of interventions  
8 reach as many people beyond the study participants as possible. Participation of CAs in the intervention  
9 resulted in positive effects on their immediate network members' HIV knowledge regardless of how the  
10 CAs responded to the intervention. While improvement in HIV knowledge may not necessarily translate to  
11 increased safe sex behavior, it can be seen as a rate-limiting step towards reducing HIV transmission since  
12 adequate knowledge is usually necessary for reducing risk behavior. Future work should examine the exact  
13 mechanisms of spillover discussed here, as that is an important clarification that could benefit future studies.  
14 The results presented herein may inform approaches for increasing participation and potentially conferring  
15 greater benefits related to spillover effects in future HIV behavioral interventions.  
16  
17  
18  
19  
20  
21  
22  
23  
24

## 25 **7. Acknowledgements**

26  
27 The authors would like to thank the study participants and staff for their time, hard work, and dedication.  
28 We would also like to thank Guy Harling for their comments on the manuscript.  
29  
30

## 31 **8. Contributorship**

32  
33 JL, MCSF, and JPO developed the research plan. JL conducted the analysis and wrote the draft of the  
34 manuscript. YL assisted in verification of analyses. All authors discussed the results and contributed to the  
35 final manuscript.  
36  
37  
38  
39  
40

## 41 **9. Compliance with ethical standards**

42  
43 **Funding:** The authors received funding from the CDC via grant TZ/UG.08.0147, the President's Emer-  
44 gency Plan for AIDS Relief - Tanzania, and the UTHealth Innovation for Cancer Prevention Research Train-  
45 ing Program Pre-doctoral Fellowship (Cancer Prevention and Research Institute of Texas grant RP160015).  
46  
47

48 **Conflict of interest:** The authors declare that they have no conflict of interest.  
49  
50  
51  
52

1  
2  
3  
4  
5       **Ethical approval:** All procedures performed in studies involving human participants were in accordance  
6 with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki  
7 declaration and its later amendments or comparable ethical standards. IRB approval was obtained from  
8 Harvard Medical School, the Muhimbili University of Health and Allied Sciences, and the National Institute  
9 for Medical Research (NIMR) in Tanzania. Particular attention was paid to eliminate the risk of inadvertent  
10 disclosure of CA HIV status to NMs.  
11  
12

13  
14  
15       **Informed consent:** Informed consent was obtained from all individual participants included in the study.

16  
17       **Disclaimer:** The content is solely the responsibility of the authors and does not necessarily represent the  
18 official views of the Cancer Prevention and Research Institute of Texas.  
19

20       **Data sharing statement:** There are no additional data available.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53



## 10. References

- [1] Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*. 2017;390(10100):1151–1210.
- [2] Crepaz N, Tungol-Ashmon MV, Higa DH, Vosburgh W, Mullins MM, Barham T, et al. A systematic review of interventions for reducing HIV risk behaviors among people living with HIV in the United States, 1988-2012. *AIDS*. 2014 mar;28(5):633–56. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24983541>.
- [3] Arnold EA, Kegeles SM, Pollack LM, Neilands TB, Cornwell SM, Stewart WR, et al. A Randomized Controlled Trial to Reduce HIV-Related Risk in African American Men Who Have Sex with Men and Women: the Bruthas Project. *Prevention Science*. 2018;p. 1–11.
- [4] Garcia de Olalla P, Knobel H, Carmona A, Guelar A, Lopez-Colomes J, Cayla J. Impact of adherence and highly active retroviral therapy on survival in HIV-infected patients. *Journal of Acquired Immune Deiciency Syndromes*. 2002;30(1):105–10.
- [5] Joint United Nations Programme on HIV/AIDS and Joint United Nations Programme on HIV/Aids and others. 90-90-90: an ambitious treatment target to help end the AIDS epidemic. Geneva: UNAIDS. 2014;.
- [6] Fisher JD, Fisher WA. Changing AIDS-risk behavior. *Psychological bulletin*. 1992;111(3):455.
- [7] McNeill C, George N, Glover R. An Evaluation of Sisters Informing Healing Living Empowering: Increasing HIV Knowledge Among African American Adolescent Females Using an Evidence-Based HIV Prevention Intervention. *Journal of Doctoral Nursing Practice*. 2017;10(1):4–10.
- [8] Carey MP, Braaten LS, Maisto SA, Gleason JR, Forsyth AD, Durant LE, et al. Using information, motivational enhancement, and skills training to reduce the risk of HIV infection for low-income urban women: A second randomized clinical trial. *Health Psychology*. 2000;19(1):3.

- 1  
2  
3  
4  
5 [9] Sinclair B, McConnell M, Green DP. Detecting spillover effects: Design and analysis of multilevel  
6 experiments. *American Journal of Political Science*. 2012;56(4):1055–1069.  
7  
8  
9 [10] Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *The New*  
10 *England journal of medicine*. 2007;357(4):370–379.  
11  
12 [11] Perkins JM, Subramanian SV, Christakis Na. Social networks and health: A systematic review of  
13 sociocentric network studies in low- and middle-income countries. *Social science & medicine*. 2014  
14 aug;125(1):60–78. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25442969>.  
15  
16  
17 [12] Coleman J, Katz E, Menzel H. The Diffusion of an Innovation among Physicians. *Sociometry*.  
18 1957;20(4):253–270.  
19  
20  
21 [13] VanderWeele TJ. Inference for influence over multiple degrees of separation on a social  
22 network. *Statistics in medicine*. 2013 feb;32(4):591–6; discussion 597–9. Available from:  
23 <http://www.ncbi.nlm.nih.gov/pubmed/23341081>.  
24  
25  
26 [14] Lienert J, Marcum CS, Finney J, Reed-Tsochas F, Koehly L. Social influence on 5-year survival in a  
27 longitudinal chemotherapy ward co-presence network. *Network Science*. 2017;5(3):308–327.  
28  
29  
30 [15] Hughes JJ. Paying injection drug users to educate and recruit their peers: why participant-driven  
31 interventions are an ethical public health model. *Quality management in health care*. 1999;7(4):4–12.  
32  
33  
34 [16] Ssali S, Wagner G, Tumwine C, Nannungi A, Green H. HIV Clients as Agents for Prevention: A Social  
35 Network Solution. *AIDS research and treatment*. 2012;2012:815823.  
36  
37  
38 [17] Cislighi B, Denny EK, Cisse' M, Gueye P, Shrestha B, Shrestha PN, et al. Changing Social Norms:  
39 the Importance of “Organized Diffusion” for Scaling Up Community Health Promotion and Women  
40 Empowerment Interventions. *Prevention Science*. 2019;p. 1–11.  
41  
42  
43 [18] Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden popula-  
44 tions for HIV surveillance. *Aids*. 2005;19:S67–S72.  
45  
46  
47 [19] Smith R, Rossetto K, Peterson BL. A meta-analysis of disclosure of one’s HIV-positive status, stigma  
48 and social support. *AIDS care*. 2008;20(10):1266–1275.  
49  
50  
51  
52

- 1  
2  
3  
4  
5 [20] Watters JK, Biernacki P. Targeted sampling: options for the study of hidden populations. *Social*  
6 *problems*. 1989;36(4):416–430.  
7  
8  
9 [21] Benjamin-chung J, Abedin J, Berger D, Clark A, Jimenez V, Konagaya E, et al. Spillover effects on  
10 *health outcomes in low- and middle-income countries: a systematic review*. *International Journal of*  
11 *Epidemiology*. 2017;(September):1–26.  
12  
13  
14 [22] Neema S, Atuyambe LM, Otolok-Tanga B, Twijukye C, Kambugu A, Thayer L, et al. Using a clinic  
15 *based creativity initiative to reduce HIV related stigma at the Infectious Diseases Institute, Mulago*  
16 *National Referral Hospital, Uganda*. *African health sciences*. 2012;12(2):231–239.  
17  
18  
19 [23] Smith Fawzi MC, Siril H, Li N, McAdam K, Ainebyona D, McAdam E, et al. Agents of Change:  
20 *stepped wedge randomized controlled trial of the NAMWEZA intervention with people living with*  
21 *HIV and members of their social networks*. *BMJ Global Health*. 2019;.  
22  
23  
24 [24] Kaaya S, Arthur G, Siril H, Lienert J, Li N, McAdam K, et al. Leadership of PLWH as Agents of  
25 *Change: comparing HIV-related risk behavior of people attending ART clinics in Dar es Salaam with*  
26 *members of their social networks*. In: *20th International AIDS Conference; 2014.*.  
27  
28  
29 [25] Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC medical research*  
30 *methodology*. 2006;6(1):54.  
31  
32  
33 [26] McCarthy B. Appreciative Inquiry: an alternative to behaviour management. *Dementia*.  
34 *2017;16(2):249–253*.  
35  
36  
37 [27] Zou J, Yamanaka Y, John M, Watt M, Ostermann J, Thielman N. Religion and HIV in Tanzania:  
38 *influence of religious beliefs on HIV stigma, disclosure, and treatment attitudes*. *BMC public health*.  
39 *2009;9(1):75*.  
40  
41  
42 [28] Carey MP, Schroder KE. Development and psychometric evaluation of the brief HIV Knowledge  
43 *Questionnaire*. *AIDS education and prevention*. 2002;14(2):172–182.  
44  
45  
46 [29] Haile BJ, Chambers JW, Garrison JL. Correlates of HIV knowledge and testing: Results of a 2003  
47 *South African survey*. *Journal of Black Studies*. 2007;38(2):194–208.  
48  
49  
50  
51  
52

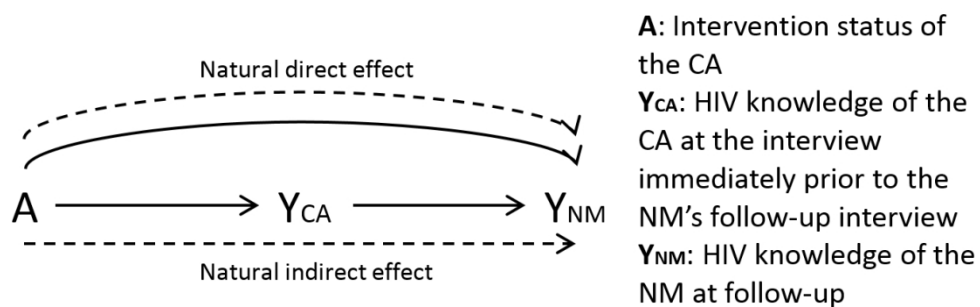
- 1  
2  
3  
4  
5 [30] Aarø LE, Breivik K, Klepp KI, Kaaya S, Onya HE, Wubs A, et al. An HIV/AIDS knowledge scale for  
6 adolescents: item response theory analyses based on data from a study in South Africa and Tanzania.  
7 Health education research. 2011;26(2):212–224.  
8  
9  
10  
11 [31] Vanderweele TJ, Hong G, Jones SM, Brown JL. Mediation and spillover effects in group-randomized  
12 trials: a case study of the 4Rs educational intervention. Journal of the American Statistical Association.  
13 2013 jun;108(502):469–482. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23997375>.  
14  
15  
16 [32] Forastiere L, Mealli F, Vanderweele TJ. Identification and Estimation of Causal Mechanisms in  
17 Clustered Encouragement Designs: Disentangling Bed Nets Using Bayesian Principal Stratification  
18 Identification and Estimation of Causal Mechanisms in Clustered Encouragement Designs : Disentan-  
19 gling Be. Journal of the American Statistical Association. 2016;111(514):510–525. Available from:  
20 <http://dx.doi.org/10.1080/01621459.2015.1125788>.  
21  
22  
23 [33] Shalizi CR, Thomas AC. Homophily and contagion are generically confounded in observational social  
24 network studies. SMR. 2012;40(2):211–39.  
25  
26  
27 [34] De La Haye K, Robins G, Mohr P, Wilson C. Homophily and contagion as explanations for weight  
28 similarities among adolescent friends. Journal of Adolescent Health. 2011;49(4):421–427.  
29  
30  
31 [35] Latkin CA, Davey-Rothwell MA, Knowlton AR, Alexander KA, Williams CT, Boodram B. Social  
32 network approaches to recruitment, HIV prevention, medical care, and medication adherence. Journal  
33 of Acquired Immune Deficiency Syndromes. 2013;63:1–9.  
34  
35  
36 [36] Galea S, Tracy M. Participation rates in epidemiologic studies. Annals of epidemiology.  
37 2007;17(9):643–653.  
38  
39  
40 [37] Langford CPH, Bowsher J, Maloney JP, Lillis PP. Social support: a conceptual analysis. Journal of  
41 advanced nursing. 1997;25(1):95–100.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**11. Figure Captions**

Fig 1. Schematic of natural direct effect (NDE) and natural indirect effect (NIE).

For peer review only



## STROBE (Strengthening The Reporting of OBServational Studies in Epidemiology) Checklist

A checklist of items that should be included in reports of observational studies. You must report the page number in your manuscript where you consider each of the items listed in this checklist. If you have not included this information, either revise your manuscript accordingly before submitting or note N/A.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).

Section and Item	Item No.	Recommendation	Reported on Page No.
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	
<b>Introduction</b>			
Background/Rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
<b>Methods</b>			
Study Design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed  <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	

Section and Item	Item No.	Recommendation	Reported on Page No.
Data Sources/ Measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study Size	10	Explain how the study size was arrived at	
Quantitative Variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	
Statistical Methods	12	(a) Describe all statistical methods, including those used to control for confounding	
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive Data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome Data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	



Section and Item	Item No.	Recommendation	Reported on Page No.
Main Results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other Analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key Results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
<b>Other Information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Once you have completed this checklist, please save a copy and upload it as part of your submission. DO NOT include this checklist as part of the main manuscript document. It must be uploaded as a separate file.**

# BMJ Open

## Network Spillover Effects and Follow-Up Correlates in a HIV Prevention Intervention in Tanzania

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-033759.R1
Article Type:	Original research
Date Submitted by the Author:	02-Jan-2020
Complete List of Authors:	Lienert, Jeffrey; University of Pennsylvania Perelman School of Medicine, Center for Health Care Innovation; Philadelphia VAMC, CHERP Fawzi, Mary C; Harvard Medical School, Global Health and Social Medicine McAdam, Keith; London School of Hygiene and Tropical Medicine, Epidemiology Kaaya, Sylvia; Muhimbili University of Health and Allied Sciences, School of Medicine Liu, Yuanyuan Todd, Jim; London School of Hygiene and Tropical Medicine, Population Health Andrew, Irene; Management and Development for Health, Onnela, Jukka Pekka; Harvard T.H. Chan School of Public Health, Biostatistics
<b>Primary Subject Heading</b>:	HIV/AIDS
Secondary Subject Heading:	Global health, Public health
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, STATISTICS & RESEARCH METHODS

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

## Network Spillover Effects and Follow-Up Correlates in a HIV Prevention Intervention in Tanzania

Jeffrey Lienert<sup>a,b,j</sup>, Mary C. Smith Fawzi<sup>c</sup>, Keith McAdam<sup>d</sup>, Sylvia Kaaya<sup>e</sup>, Yuanyuan Liu<sup>i</sup>, Jim Todd<sup>g</sup>, Irene Andrew Lema<sup>f</sup>, Jukka-Pekka Onnela<sup>h</sup>

<sup>a</sup>Perelman School of Medicine, University of Pennsylvania <sup>b</sup>Center for Health Equity Research and Promotion, Philadelphia VA <sup>c</sup>Department of Global Health and Social Medicine, Harvard Medical School <sup>d</sup>Department of Clinical Tropical Medicine, London School of Hygiene and Tropical Medicine <sup>e</sup>Department of Psychiatry and Mental Health, Muhimbili University of Health and Allied Sciences <sup>f</sup>Management and Development for Health <sup>g</sup>Department of Population Health, London School of Hygiene and Tropical Medicine <sup>h</sup>Department of Biostatistics, Harvard T. H. Chan School of Public Health <sup>i</sup>University of Texas School of Public Health <sup>j</sup>Corresponding Author: Jeffrey Lienert, [jeffrey.lienert@penntmedicine.upenn.edu](mailto:jeffrey.lienert@penntmedicine.upenn.edu)

## Abstract

**Objectives:** We aim to describe the social network members of participants of a behavioral intervention, and examine how the effects of the intervention may spillover among network members.

**Design:** Secondary analysis of a step-wedge randomized controlled trial.

**Setting:** Change Agents (CAs) were recruited from waiting rooms of HIV treatment facilities in Dar es Salaam, Tanzania, and their Network Members (NMs) were recruited directly by CAs.

**Participants:** We enrolled 662 CAs in an HIV behavioral intervention. They, along with 710 of their NMs, completed baseline and follow-up interviews from 2011 to 2013.

**Primary and Secondary Outcomes:** The primary outcome of this study was change in NMs' HIV knowledge, and the secondary outcome was whether the NM was lost to follow-up.

**Results:** At baseline, many characteristics were different between NMs and CAs. We found a number of NM characteristics significantly associated with follow-up of NMs, particularly female gender (OR=1.64, 95% CI: 1.02,2.63) and HIV knowledge (OR=20.0, 95% CI: 3.70,125); only one CA variable was significantly associated with NM follow-up: having a private source of water (OR=2.17, 95% CI: 1.33,3.57). The 14.2% increase in NMs' HIV knowledge was largely due to CAs feeling empowered to pass on prior knowledge, rather than transmitting new knowledge to their NMs.

**Conclusions:** Characteristics of social network members of PLH may play a role in study retention. Additionally, the HIV knowledge of these NMs increased largely as a function of CA participation in the intervention, indicating that intervening in highly-connected individuals may maximize benefits to the potential population for whom spillover can occur.

*Keywords:*

mediation, natural direct effect, natural indirect effect, follow-up, retention, hidden population

## 1. Strengths and Limitations of This Study

- The recruitment method, focused on social network members of persons living with HIV, allowed us to access an at-risk population typically not easily-accessible via other means, such as direct approach by the study team rather than someone known to the participant.
- The study design of the trial in which this study is nested was an ideal situation in which to employ a mediation analysis that informs our understanding of how this and similar intervention effects may spillover in a population.
- Although the high dropout rate of social network members meant we could examine factors leading to dropout, estimation of the outcome will be biased if those who dropped out were different from those who did not.

## 2. Introduction

Despite recent decreases in mortality, HIV/AIDS is currently the ninth leading cause of years of life lost (YLL) globally [1]. Research into eliminating HIV has focused on two fronts: biological and behavioral, which combined have decreased the YLL due to HIV by an estimated 51% [1]. Behavioral interventions seek to curb transmission of HIV through reduction of risk behavior including safer sexual practices and reducing injection drug use [2, 3]. Behavioral change also impacts the effectiveness of the biological methods, given that lower adherence is known to often compromise their effectiveness [4]. However, neither biological nor behavioral methods are fully effective without the knowledge of these treatments and preventative behaviors. It is therefore important to increase HIV-related knowledge as a necessary, but not sufficient, step towards achieving the 90-90-90 treatment goals and ultimately eliminate HIV [5].

The importance of adequate HIV knowledge has been recognized for a period of time, as it is often a prerequisite for behavioral change [6]. It should be noted that increased knowledge does not necessarily translate into behavior change, as other factors such as motivation and skills also play a role [6]. Despite this, researchers conducted trials among persons living with HIV (PLH) which led to increased HIV knowledge [7, 8]. These interventions were group discussions and teaching led by nurses, and motivational interview-

1  
2  
3  
4  
5 ing, respectively. Researchers also found that the intervention reduced HIV-risky behaviors  
6 concomitantly with an increase in HIV knowledge [8].  
7

8 Knowledge gained by participants in these trials can also be freely shared with members of their social  
9 network, in what is known as a spillover effect [9]. Specifically, a spillover effect (also known as an in-  
10 direct or disseminated effect) is one person's exposure affecting another's outcome [10]. In the context of a  
11 social network spillover intervention, it means a change to someone's behavior who did not receive the  
12 intervention because they were socially connected to someone who did receive the intervention. This is dis-  
13 tinct from what is sometimes called behavioral spillover, where change's in a person's behavior affects other  
14 behaviors of that same person [11]. For injection drug users, interventions have shown spillover effects of  
15 HIV prevention education, and subsequent reduced rates of risky behaviors [12]. What remains unknown,  
16 however, is whether or not a spillover effect exists for HIV knowledge during and after an intervention in  
17 other populations, particularly sub-Saharan Africa.  
18  
19  
20  
21  
22  
23  
24

25 New knowledge can come from a variety of sources, one of the most important of which is a person's  
26 so- cial network [13, 14]. Social networks are of particular import because new knowledge can lead to  
27 cascades of behavior change, where people subsequently educate those in their social network, in what is  
28 known as social influence [15, 16, 17]. This has been directly examined in participant-driven interventions,  
29 where initially-recruited participants directly educate members of their social network [18]. Characteristics  
30 such as knowing the HIV status of network members has been shown to be the most important predictor of  
31 engaging in prevention advocacy [19]. Work on diffusion through social networks has shown that spreading  
32 inter- vention effects beyond the initial study population can improve the cost-effectiveness of these  
33 interventions [20]. These findings imply that certain aspects of knowledge or behavior may spread more or  
34 less efficiently through networks comprising individuals with specific characteristics, which may need to be  
35 accounted for in network interventions. For instance, networks comprising many at-risk individuals who are  
36 HIV-negative may not be as receptive to change in behaviors as networks comprising a mix of those with  
37 and without HIV. Additionally, the networks of PLH are often difficult networks to ascertain [21], due to  
38 the continued stigma of HIV and AIDS in many settings [22]. Because of this, if a PLH or person at-risk of  
39 acquiring HIV does not want to participate in an investigator-initiated intervention, there is little recourse  
40 other than information transmitted via social networks, or targeted sampling techniques which are not always  
41 effective  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

1  
2  
3  
4  
5 [23]. This is particularly important in low- and middle-income countries, as a recent systematic review  
6 found only 54 studies researching spillover effects [24]. Therefore, understanding exactly how information  
7 spreads from participants in an intervention to members of their social network, who may be largely  
8 inaccessible via other means, is important for reaching the greatest number of people about HIV prevention.  
9 Understanding what makes these persons different from those who receive the intervention themselves is  
10 important, as it may point to ways in which to increase enrollment of these populations.  
11  
12

13  
14  
15 Based on the above gaps in the literature, we conducted a study on network members of PLH enrolled  
16 in the larger NAMWEZA intervention [25]. The trial recruited PLH to serve as Change Agents (CAs) and  
17 to reach out to their social network members (NMs) about knowledge of HIV and safer sexual practices  
18 [26]. Our goals in this study were threefold: 1) to understand how the NMs differed from their CAs, 2)  
19 understand correlates of dropout for the NMs, and 3) to understand by what method HIV knowledge  
20 transferred to the NMs from the CAs. Understanding how the information and behaviors are shared within  
21 social networks will allow HIV researchers and others to take advantage of this knowledge and improve  
22 upon prevention interventions in the future.  
23  
24  
25  
26  
27  
28

### 29 30 **3. Methods**

#### 31 32 *3.1. Study Population*

33  
34 We analyzed social network data from the Agents of Change trial [27], which was a stepped-wedge  
35 randomized controlled trial that enrolled PLH to become Change Agents (CAs) by informing members of  
36 their social network (NMs) about knowledge of HIV and safer sexual practices [28]. Here, we define CAs  
37 based on the *potential* for PLH to become so - by self-selecting into the study, PLH identify themselves as  
38 potential CAs, which we aim to foster through the intervention.  
39  
40  
41

42 CAs were recruited from the waiting rooms of HIV care and treatment centers in Dar es Salaam, Tanza-  
43 nia. Participants were given a baseline questionnaire and were randomized to one of three waves in which  
44 to receive the intervention. At baseline, participants were also asked to recruit up to three members of their  
45 social networks who they felt were at particularly high risk of contracting or spreading HIV. These network  
46 members (NMs) could be either HIV positive or negative, and they were given a baseline survey. The NM  
47 was only aware the CA was a participant in the intervention if the CA shared this information with them,  
48  
49  
50  
51  
52



1  
2  
3  
4  
5 which many did not due to HIV-related stigma [29]. Each CA therefore formed a CAN-NM dyad with each  
6 NM they recruited, and if they recruited more than one NM, formed an egocentric network with multiple  
7 dyads.  
8  
9

10 For CAs, the intervention comprised 10 weekly structured sessions aimed at empowering PLH to be-  
11 come HIV prevention change agents in their communities. The sessions aimed to increase psychosocial and  
12 communication skills through an Appreciative Inquiry approach [30]. Within one month of each wave of the  
13 intervention, CAs were given follow-up surveys. The interventions lasted from November 2010 to January  
14 2014.  
15  
16  
17

18 For NMs, study staff did not offer any direction intervention; the CAs provided information directly to  
19 their NMs. Rather, their intervention status flowed from the intervention status of their CAs. Due to opera-  
20 tional difficulties, each NM was surveyed one time during the study. We use this interview of NMs as the  
21 division between “exposed” and “unexposed.” At the time of an NM’s follow-up interview, not all of their  
22 respective CAs had undergone the intervention. Therefore, the NMs were divided into “exposed” (N=381)  
23 and “unexposed” (N=329) groups based on whether their respective CA had completed their intervention at  
24 the time of the NM’s first follow-up interview. In this way, we were able to assess the longitudinal spillover  
25 effect of the intervention net of temporal or geographical trends.  
26  
27  
28  
29  
30

31 Inclusion criteria for the follow-up analyses were for CAs who completed at least one follow-up inter-  
32 view and had at least one NM who completed at least one follow-up questionnaire. In this way, outcomes  
33 could be computed for the CA and NM of each dyad. NMs approached by a CA, but who did not participate  
34 in the study were not recorded since it was not feasible to obtain this information from study participants  
35 themselves. In sum, our sample comprises 662 CAs and 710 NMs, meaning each CA recruited 1.07 NMs  
36 on average out of a possible 3.  
37  
38  
39  
40

41 During this study, there was little loss-to-follow-up among the CAs (< 10%), but much higher among  
42 the NMs [31].  
43  
44  
45

### 46 3.2. *HIV knowledge*

47 To assess HIV knowledge of CAs and NMs, we used the brief HIV knowledge questionnaire [32]. This  
48 scale comprises 18 questions, which focus on an individual’s knowledge of how HIV can spread and other  
49  
50  
51  
52

1  
2  
3  
4  
5 characteristics of the virus and AIDS. This knowledge is crucial for an individual's subsequent safe-sex  
6 practices to reduce the risk of transmission. The original population comprised three different groups: two  
7 groups of low-income women, and one of women and men receiving psychiatric treatment. These popula-  
8 tions had a Chronbach's alpha of 0.78. This instrument has been used previously in sub-Saharan Africa and  
9 has demonstrated reasonable reliability of 0.75 Cronbach's alpha in South Africa among a convenience sam-  
10 ple of 429 members of the African Methodist Episcopal church [33]. It has also been translated to Swahili,  
11 with only minor differential item functioning [34]. This indicates that the measure performs adequately in  
12 other, similar populations, relevant to this work. Here, the Cronbach's alpha was 0.71. Because the aver-  
13 age baseline score of HIV knowledge was 0.80, with a preponderance of scores of 1.00, the main indicator  
14 of knowledge was not normally distributed, and therefore a linear regression was not ideal. We therefore  
15 summarize this measure as "Complete HIV Knowledge", a dichotomous variable for whether the participant  
16 correctly answered all questions.  
17  
18  
19  
20  
21  
22  
23  
24  
25

### 26 3.3. *Demographic and contextual variables*

27  
28 In addition to HIV knowledge, we are interested in a set of demographic and contextual variables that  
29 may help explain some of the trends observed. In terms of demographic variables, we include age, sex,  
30 employment status, marital status, and self-identified HIV status. We also include education, which we di-  
31 chotomize at 7 years or more - a level commensurate with elementary education. We selected this cutoff  
32 because it coincides with the millennium development goal (MDG) of increasing primary education com-  
33 pletion [35].  
34  
35  
36  
37

38 Contextually, we included two additional variables: having a private source of water, and the number  
39 of persons sleeping in the participant's home. The first was also based on a MDG, and indicates participants  
40 with access to safe drinking water [35]. This is a proxy for the economic security of the participant. The  
41 number of persons sleeping in the participant's home is also a non-monetary indicator of their material and  
42 social resources [36]. These variables combined give a more thorough picture of the participant's economic  
43 status than employment alone.  
44  
45  
46  
47  
48  
49  
50  
51  
52

### 3.4. Statistical Analysis

The first analysis was to assess homophily, defined as the extent to which CAs and NMs were similar to one another in terms of a number of sociodemographic characteristics and HIV-related risk factors. Because CAs could have more than one NM, we assessed statistical significance of homophily using univariate intercept-only Generalized Estimating Equations (GEE) with a Normal or Bernoulli distribution for continuous and dichotomous variables, respectively. We used exchangeable working correlation structures in these models. Analyses were run using R v3.1.1.

To accomplish our second aim of understanding what was associated with NMs completing all their follow-up interviews, we fit a GEE to determine predictors of follow-up, clustered at the level of the CA. In this regression, we use all the variables listed above, as well as whether the CA was lost to follow-up (LTFU). In order to examine the association between these same variables and the time to follow-up, we also employed a Cox regression. In the Cox regression, start of follow-up was defined as the time at which the NM completed their baseline interview. The outcome was loss-to-follow-up. NMs who were lost to follow-up were censored at 3 months after their last interview, the point at which the study-defined criteria for LTFU was met [37].

Finally, since the trial showed beneficial effects on the HIV knowledge of the NMs following the intervention [26], we aimed to elucidate exactly what caused the HIV knowledge of NMs to increase - either CAs gaining knowledge through the intervention and passing it on, or the CAs being empowered by the intervention to pass on existing knowledge. As shown by VanderWeele et al (2015), social network spillover effects can be broken down into direct and indirect effects in the case of dyadic relationships (Figure 1) [38]. This method has since been used for novel evaluations of spillover effects [39]. Although previous studies showed that this same type of analysis cannot be done on full network data, the data in this study consisted of only paired individuals, the CAs and their NMs, so in this case the analysis does not result in biased estimates [40, 41].

One assumption of this analysis is that each pair is totally independent, which is violated here, if a CA recruited more than one NM - the multiple CA-NM dyads including 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.

1  
2  
3  
4  
5 The method parses social influence into direct and indirect effects. The natural indirect effect (NIE) is  
6 the effect by which the intervention changes the CA's HIV knowledge, which then changes the NM's HIV  
7 knowledge. The NIE is similar to the effect observed in many participant-driven interventions: an initial  
8 participant receives the intervention, increasing their knowledge, and they subsequently pass their increased  
9 knowledge to members of their social network [18]. The natural direct effect (NDE) is the effect of receiving  
10 the intervention has on an NMs outcome, irrespective of the CA's outcome (in this case HIV knowledge).  
11 For instance, this could occur if CAs begin with good knowledge of HIV, and the intervention empowers  
12 them to convey knowledge they already had to their NMs. Although the intervention does not increase their  
13 HIV knowledge, it is still useful to the CAs, as it empowers them to become CAs in the first place. In order  
14 to estimate these effects, the published SAS macro developed by VanderWeele and colleagues was used,  
15 and analyses were run using SAS v9.2 (SAS Institute, Cary, NC) [42]. The data are not publicly available  
16 due to the sensitive nature of HIV infection status.  
17  
18  
19  
20  
21  
22  
23  
24  
25

### 26 3.5. Patient and Public Involvement

27  
28 Patients were not involved in the design of this study. As part of the recruitment process, patients were  
29 instructed to recruit members of their social networks who they felt were at high risk of HIV infection - in  
30 this way participants could guide the recruitment of NMs to those most at-risk. We currently have no plans  
31 to disseminate the results of this study to participants.  
32  
33  
34  
35

## 36 4. Results

37  
38 The *NAMWEZA* study recruited 662 Change agents (CA), of whom 453 (68.4%) completed at least one  
39 follow-up questionnaire. These 662 CAs in turn recruited 710 network members (NM) who took the baseline  
40 questionnaire. Of the 710 NMs, 449 (63.2%) also completed a follow-up questionnaire (Table 1). At baseline,  
41 CAs were on average older than their NMs. More of the NMs were employed than their CAs (69.3% vs.  
42 55.0%,  $p < 0.0001$ ), but were less likely to have at least 7 years of education (52.0% vs. 52.3%,  $p < 0.0001$ ).  
43 Only 12.3% of NMs were HIV-positive, compared to *all* CAs ( $p < 0.0001$ ). Complete data was obtained at  
44 baseline for all CAs and NMs.  
45  
46  
47  
48  
49  
50  
51  
52

Characteristic	Number of NMs (%) or Mean (SD) (N=710)	Number of CAs (%) or Mean (SD) (N=662)	P-value
Age	33.0 (11.1)	38.9 (9.7)	<0.0001
Female	380 (53.7%)	349 (53.9%)	0.956
Employed	490 (69.3%)	356 (55.0%)	<0.0001
At least 7 years education	369 (52.0%)	584 (82.3%)	<0.0001
Complete HIV knowledge	638 (89.9%)	598 (90.4%)	0.78
Persons sleeping in home	5.00 (5.77)	3.87 (3.72)	<0.0001
Married	373 (52.7%)	338 (51.1%)	0.61
HIV Positive	87 (12.3%)	662 (100%)	<0.0001
Private source of water	309 (43.7%)	263 (39.7%)	0.20

Table 1: Demographic characteristics at baseline, with the results of an intercept-only GEE for differences. Specifically, for each CA-NM dyad, either a difference (for continuous variables) or concordance (for dichotomous variables) is calculated. These are then used as an outcome in an intercept-only GEE clustering on CA, and the p-value of the intercept is shown.

Logistic regression showed that characteristics of both the CAs and the NMs significantly predicted loss to follow-up (Table 2). The NM being female (OR=1.64), having complete HIV knowledge (OR=20), being employed (OR=1.5), and being married (OR=1.72) were all significant predictors of increased odds of being followed-up. Each additional person sleeping in the home of the NM per room used for sleeping reduced the odds of follow-up (OR=0.81, 95% CI: 0.67,0.98) as did the NM living with HIV (OR=0.42, 95% CI: 0.18,0.99). CA having a private water source was significantly associated with increased odds of the NM being followed-up (OR=2.17, 95% CI: 1.33,3.57), even after controlling for the NM having a private source of water. This was the only CA-specific variable that significantly predicted an NM's follow-up.

The Cox proportional hazard model showed similar results to the logistic regression, but with fewer significant results. Only the NM being married (HR=1.28) and the NM living with HIV (HR=0.71) significantly predicted the time-to-follow-up of the NMs. Additionally, all of the hazard ratios were closer to the null relative to the corresponding odds ratios.

We assess the mechanism of spillover effects on HIV knowledge, the main effect of the intervention on the NMs. The analysis showed that NMs who were exposed to intervention via their CA had an increase in HIV knowledge of 12.9% on average (95% CI: 0.06, 0.20) from a baseline percentage of 78% [26]. This was broken down into a natural indirect effect (NIE) of 0.6% (95% CI: -0.06%,2.0%), which is the effect the intervention had via the CA's HIV knowledge, and a natural direct effect (NDE) of 12.3% (95% CI:

Characteristic (N=459)	Adjusted OR (95% CI)	Adjusted HR (95% CI)
Characteristics of NMs		
Gender (Female)	1.64* (1.02,2.63)	1.18 (0.94,1.50)
Difference in age of NM and CA (per year)	1.01 (0.99,1.03)	1.01 (0.995,1.02)
Age of NM	0.98 (0.95,1.02)	0.99 (0.97,1.04)
Complete HIV knowledge (vs. none)	20* (3.70,125)	2.20 (0.97,5.01)
Employed	1.5* (1.04,2.5)	1.15 (0.89,1.50)
Each additional person sleeping in home per room used for sleeping	0.81* (0.67,0.98)	0.92 (0.83,1.01)
Married	1.72* (1.04,2.86)	1.28* (1.01,1.64)
Living with HIV	0.42* (0.18,0.99)	0.71* (0.51,0.99)
Having a private source of water	0.86 (0.54,1.37)	0.97 (0.77,1.22)
Characteristics of CAs		
Gender (Female)	1.27 (0.76,2.08)	1.07 (0.84,1.37)
Complete HIV knowledge (vs. none)	0.36 (0.07,2.04)	0.64 (0.29,1.43)
Having a private source of water	2.17* (1.33,3.57)	0.97 (0.77,1.22)
Being employed	1.54 (0.95,2.50)	1.14 (0.90,1.43)
Being married	1.25 (0.77,2.04)	1.11 (0.88,1.43)
CA Lost to follow-up	1.06 (0.84,1.33)	1.02 (0.72,1.43)

Table 2: Results of multivariate logistic regression and Cox-proportional hazard models on the dichotomous outcomes of whether the NM was followed-up, and the continuous outcome of time-to-follow-up, respectively. “\*” indicates significant at the  $p < 0.05$  level.

6.1%,19.3%), which is the effect the CA’s participation had on the NM’s HIV knowledge, irrespective of the CA’s HIV knowledge. In other words, of the 12.9% increase in NM’s having complete HIV knowledge, 12.3% occurred without a concomitant increase in their CA’s HIV knowledge - their HIV knowledge increase was not mediated by the increase in HIV knowledge of their CA. This did not change when we used only one NM per CA.

## Discussion

Based on results from the NAMWEZA trial in Tanzania, we have shown a number of novel findings regarding the network members of HIV intervention participants, correlates of their retention in the study, and the evidence for spillover of HIV knowledge from CAs to NMs. These findings can inform the design of interventions in the future to maximally enroll and retain participants and ensure that maximal correct information transmitted from the study participants to members of their social networks.

We found many significant differences between CAs and the NMs they recruited. On average, CAs were older, less likely to be employed, and more educated. All of this suggests that by having CAs recruit from

1  
2  
3  
4  
5 their social network, we were able to recruit a set of social network members very different from those found  
6 in the waiting rooms in HIV treatment clinics (unsurprising, given that many NMs were not HIV-positive)  
7 [9]. This may in part be due to the fact that when recruiting NMs, CAs were instructed to recruit those who  
8 may be at particular risk of contracting or spreading HIV. Therefore, CAs likely did not pick random  
9 members of their social network, but those who were at high risk. This finding corroborates the claim by  
10 Latkin et al. (2013) that sampling via enumeration of network members by initial participants recruits a more  
11 diverse sample [43]. Additionally, the lower likelihood of employment of the CAs indicates they may have  
12 had more time for participating in the intervention [44]. This means that future studies may need to tailor  
13 their interventions to work with the schedules of employed persons to ensure participation.  
14  
15  
16  
17  
18  
19

20 Although the study design potentially accessed a separate slice of the population than other methods  
21 vis-a'-vis the population of those at risk for contracting HIV, it may have been at least partially responsible  
22 for the large loss of follow-up of the NMs. It is important to note that more NM characteristics than CA  
23 characteristics were significantly predictive of follow-up in both the logistic and Cox proportional hazards  
24 models. We also note that less HIV knowledge and living with HIV predicted greater loss to follow-up  
25 among NMs, which means that those who might have benefited most from spillover of the intervention were  
26 more likely to discontinue their involvement. This does not mean that they did not receive any spillover,  
27 only that it was not recorded. Our estimate of the magnitude of the spillover may therefore be biased towards  
28 the null. This is also problematic more generally for interventions of this nature as the very people the  
29 intervention aims to benefit may not stay with the program.  
30  
31  
32  
33  
34  
35

36 The HIV knowledge gain experienced by the NMs was largely due to the NDE - i.e. knowledge spilled-  
37 over as a result of the CAs participating in NAMWEZA, not because the CAs' HIV knowledge increased,  
38 and they passed this new knowledge to their NM(s). This is most likely because CAs had a high average  
39 HIV knowledge score at baseline (80%), so there was less room for improvements in knowledge. Following  
40 the intervention, CAs' likelihood of complete HIV knowledge did not significantly increase. Therefore, what  
41 prompted the increased knowledge of the NMs was the CA becoming empowered through the intervention  
42 to pass on their existing knowledge to their NM [45].  
43  
44  
45  
46  
47

48 This finding is important for future interventions: spillover effects of this intervention will likely carry  
49 over only one degree of separation from the CAs (opposed to spreading indefinitely in a snowball effect),  
50  
51  
52

1  
2  
3  
4  
5 as the most important thing in transmitting HIV knowledge is receiving the intervention, which only CAs  
6 receive. Because NMs do not actually receive the intervention, but only the contact with their CA, it is  
7 unlikely that they would feel empowered to become CAs themselves (particularly because they will not  
8 receive the *NAMWEZA* sessions), thus limiting the spread of the intervention. For CAs to increase the HIV  
9 knowledge of their NMs, they would only need to become empowered to share their information. However,  
10 for the NMs to increase the HIV knowledge of their own NMs, they would need to both increase their HIV  
11 knowledge, and become empowered to share it. This is less likely to happen than just becoming empowered,  
12 and so it is unlikely, though not impossible, that this effect would continue to spread in the population. This  
13 may give insight into how to design interventions in the future; if one wants to maximize the number of  
14 people who benefit, choosing CAs who form many bridging ties in the community would maximize the  
15 potential number of links by which spillover can occur [46]. Alternatively, interventions can be designed to  
16 be self-propagating; if CAs are empowered to deliver the intervention to others, changing their own NMs  
17 into future CAs, the effects seen here could spillover continuously.

18  
19  
20  
21  
22  
23  
24  
25  
26  
27 Our study is not without limitations. First, there was considerable loss to follow-up of the NMs. Al-  
28 though this actually informed our analysis of the correlates of loss to follow-up, it meant that our final  
29 analysis may have biased our results. Even though the exposure was randomized, the loss to follow-up can  
30 result in selection bias if the NMs who left the study were systematically different from those who remained.  
31 As we show, the NMs who dropped out were those who would have benefited the most from the intervention  
32 because they were more vulnerable and at higher risk than those who did not drop out. Because they are  
33 likely to have greater benefits from any spillover effects, we would expect this non-random loss to follow-up  
34 to result in an underestimate of the impact of the intervention in the network members. Second, our data did  
35 not perfectly fit the requirements of the causal mediation analysis: they were not entirely independent, since  
36 multiple NM-CA dyads shared a CA. However, when we randomly removed dyads until there were no  
37 repeated CAs, the results were qualitatively very similar, indicating that lack of independence did not unduly  
38 affect our results. Third, although we were able to tease apart the direct and indirect effects, we are unable  
39 to determine the mechanism of the natural direct effect; the data do not allow us to specify whether NMs  
40 increased knowledge through speaking to their knowledgeable CA, through researching HIV on their own,  
41 or some other mechanism. Future work will have to be done to examine these different pathways. Fourth, it  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51



1  
2  
3  
4  
5 is possible that there was contamination between CAs if multiple CAs knew the same NM in what is known  
6 as partial interference [47] . However, given the large population of Dar es Salaam, and the small number  
7 of NMs recruited by each CA, we do not think this is a significant problem.  
8  
9

## 10 11 **5. Conclusions**

12  
13  
14 These results have implications for the potential scale-up of the *NAMWEZA* intervention, as well as  
15 future studies and interventions that focus on behavioral interventions in social networks. First, our findings  
16 of minimal similarity between CAs and their NMs indicate that this recruitment method allows us to enroll  
17 participants from portions of the population that are not represented by the CAs alone. Understanding how  
18 this then translates to how relationally close NMs are to their CAs, and the implications for follow-up may  
19 inform strategies for retention and may increase the impact of future HIV behavioral interventions. The  
20 mediation analysis presents a compelling picture of how best to ensure the benefits of interventions reach as  
21 many people beyond the study participants as possible. Participation of CAs in the intervention resulted in  
22 positive effects on their immediate network members' HIV knowledge regardless of how the CAs responded  
23 to the intervention. While improvement in HIV knowledge may not necessarily translate to increased safe  
24 sex behavior, it can be seen as a rate-limiting step towards reducing HIV transmission since adequate  
25 knowledge is usually necessary for reducing risk behavior. Future work should examine the exact  
26 mechanisms of spillover discussed here, as that is an important clarification that could benefit future studies.  
27 Specifically, a similar setup to the study here combined with semi-structured interviews with CAs and NMs  
28 about their interactions with one another would help elucidate exactly how NMs increased their HIV  
29 knowledge (this population would likely need to comprise CAs who have disclosed their HIV sta- tus in  
30 order to prevent accidental disclosure). Future work should also examine whether increases in HIV  
31 knowledge translate to changes in behaviors putting individuals at risk of contracting HIV, particularly in  
32 Sub-Saharan Africa. The results presented herein may inform approaches for increasing participation and  
33 potentially conferring greater benefits related to spillover effects in future HIV behavioral interventions.  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

## 6. Acknowledgements

The authors would like to thank the study participants and staff for their time, hard work, and dedication. We would also like to thank Guy Harling for their comments on the manuscript.

## 7. Contributorship

JL, MCSF, and JPO designed the analysis plan and drafted the initial paper. JPL, MCSF, KM, SK, IAL and JT conducted the underlying randomized trial, and designed it to incorporate social networks data and spillover information; these same authors provided feedback on the proposed analysis plan. JL and YL conducted analyses for the paper. All authors helped revise the drafts of the paper and approved the final manuscript.

## 8. Compliance with ethical standards

**Funding:** The authors received funding from the CDC via grant TZ/UG.08.0147, the President's Emergency Plan for AIDS Relief - Tanzania, and the UTHealth Innovation for Cancer Prevention Research Training Program Pre-doctoral Fellowship (Cancer Prevention and Research Institute of Texas grant RP160015).

**Conflict of interest:** The authors declare that they have no conflict of interest.

**Ethical approval:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. IRB approval was obtained from Harvard Medical School, the Muhimbili University of Health and Allied Sciences, and the National Institute for Medical Research (NIMR) in Tanzania. Particular attention was paid to eliminate the risk of inadvertent disclosure of CA HIV status to NMs.

**Informed consent:** Informed consent was obtained from all individual participants included in the study.

**Disclaimer:** The content is solely the responsibility of the authors and does not necessarily represent the official views of the Cancer Prevention and Research Institute of Texas.

**Data sharing statement:** There are no additional data available.

## 9. References

- [1] Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*. 2017;390(10100):1151–1210.
- [2] Crepaz N, Tungol-Ashmon MV, Higa DH, Vosburgh W, Mullins MM, Barham T, et al. A systematic review of interventions for reducing HIV risk behaviors among people living with HIV in the United States, 1988-2012. *AIDS*. 2014 mar;28(5):633–56. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24983541>.
- [3] Arnold EA, Kegeles SM, Pollack LM, Neilands TB, Cornwell SM, Stewart WR, et al. A Randomized Controlled Trial to Reduce HIV-Related Risk in African American Men Who Have Sex with Men and Women: the Bruthas Project. *Prevention Science*. 2018;p. 1–11.
- [4] Garcia de Olalla P, Knobel H, Carmona A, Guelar A, Lopez-Colomes J, Cayla J. Impact of adherence and highly active retroviral therapy on survival in HIV-infected patients. *Journal of Acquired Immune Deiciency Syndromes*. 2002;30(1):105–10.
- [5] Joint United Nations Programme on HIV/AIDS and Joint United Nations Programme on HIV/Aids and others. 90-90-90: an ambitious treatment target to help end the AIDS epidemic. Geneva: UNAIDS. 2014;.
- [6] Fisher JD, Fisher WA. Changing AIDS-risk behavior. *Psychological bulletin*. 1992;111(3):455.
- [7] McNeill C, George N, Glover R. An Evaluation of Sisters Informing Healing Living Empowering: Increasing HIV Knowledge Among African American Adolescent Females Using an Evidence-Based HIV Prevention Intervention. *Journal of Doctoral Nursing Practice*. 2017;10(1):4–10.
- [8] Carey MP, Braaten LS, Maisto SA, Gleason JR, Forsyth AD, Durant LE, et al. Using information, motivational enhancement, and skills training to reduce the risk of HIV infection for low-income urban women: A second randomized clinical trial. *Health Psychology*. 2000;19(1):3.

- 1  
2  
3  
4  
5 [9] Sinclair B, McConnell M, Green DP. Detecting spillover effects: Design and analysis of multilevel  
6 experiments. *American Journal of Political Science*. 2012;56(4):1055–1069.  
7  
8  
9 [10] Benjamin-Chung J, Arnold BF, Berger D, Luby SP, Miguel E, Colford Jr JM, et al. Spillover effects in  
10 epidemiology: parameters, study designs and methodological considerations. *International journal of*  
11 *epidemiology*. 2017;47(1):332–347.  
12  
13  
14 [11] Truelove HB, Carrico AR, Weber EU, Raimi KT, Vandenberg MP. Positive and negative spillover of  
15 pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental*  
16 *Change*. 2014;29:127–138.  
17  
18  
19 [12] Ghosh D, Krishnan A, Gibson B, Brown SE, Latkin CA, Altice FL. Social network strategies to address  
20 HIV prevention and treatment continuum of care among at-risk and HIV-infected substance users: a  
21 systematic scoping review. *AIDS and Behavior*. 2017;21(4):1183–1207.  
22  
23  
24 [13] Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *The New*  
25 *England journal of medicine*. 2007;357(4):370–379.  
26  
27  
28 [14] Perkins JM, Subramanian SV, Christakis Na. Social networks and health: A systematic review of  
29 sociocentric network studies in low- and middle-income countries. *Social science & medicine*. 2014  
30 aug;125(1):60–78. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25442969>.  
31  
32  
33 [15] Coleman J, Katz E, Menzel H. The Diffusion of an Innovation among Physicians. *Sociometry*.  
34 1957;20(4):253–270.  
35  
36  
37 [16] VanderWeele TJ. Inference for influence over multiple degrees of separation on a social network.  
38 *Statistics in medicine*. 2013 feb;32(4):591–6; discussion 597–9. Available from:  
39 <http://www.ncbi.nlm.nih.gov/pubmed/23341081>.  
40  
41  
42 [17] Lienert J, Marcum CS, Finney J, Reed-Tsochas F, Koehly L. Social influence on 5-year survival in a  
43 longitudinal chemotherapy ward co-presence network. *Network Science*. 2017;5(3):308–327.  
44  
45  
46 [18] Hughes JJ. Paying injection drug users to educate and recruit their peers: why participant-driven  
47 interventions are an ethical public health model. *Quality management in health care*. 1999;7(4):4–12.  
48  
49  
50  
51  
52

- 1  
2  
3  
4  
5 [19] Ssali S, Wagner G, Tumwine C, Nannungi A, Green H. HIV Clients as Agents for Prevention: A Social  
6 Network Solution. *AIDS research and treatment*. 2012;2012:815823.  
7  
8  
9 [20] Cislighi B, Denny EK, Cisse' M, Gueye P, Shrestha B, Shrestha PN, et al. Changing Social Norms: the  
10 Importance of "Organized Diffusion" for Scaling Up Community Health Promotion and Women  
11 Empowerment Interventions. *Prevention Science*. 2019;p. 1–11.  
12  
13  
14 [21] Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden popula-  
15 tions for HIV surveillance. *Aids*.2005;19:S67–S72.  
16  
17  
18 [22] Smith R, Rossetto K, Peterson BL. A meta-analysis of disclosure of one's HIV-positive status, stigma  
19 and social support. *AIDS care*.2008;20(10):1266–1275.  
20  
21  
22 [23] Watters JK, Biernacki P. Targeted sampling: options for the study of hidden populations. *Social*  
23 *problems*. 1989;36(4):416–430.  
24  
25  
26 [24] Benjamin-chung J, Abedin J, Berger D, Clark A, Jimenez V, Konagaya E, et al. Spillover effects on  
27 health outcomes in low- and middle-income countries: a systematic review. *International Journal of*  
28 *Epidemiology*. 2017;(September):1–26.  
29  
30  
31 [25] Neema S, Atuyambe LM, Otolok-Tanga B, Twijukye C, Kambugu A, Thayer L, et al. Using a clinic  
32 based creativity initiative to reduce HIV related stigma at the Infectious Diseases Institute, Mulago  
33 National Referral Hospital, Uganda. *African health sciences*. 2012;12(2):231–239.  
34  
35  
36 [26] Smith Fawzi MC, Siril H, Li N, McAdam K, Ainebyona D, McAdam E, et al. Agents of Change:  
37 stepped wedge randomized controlled trial of the NAMWEZA intervention with people living with  
38 HIV and members of their social networks. *BMJ Global Health*. 2019;.  
39  
40  
41 [27] Kaaya S, Arthur G, Siril H, Lienert J, Li N, McAdam K, et al. Leadership of PLWH as Agents of  
42 Change: comparing HIV-related risk behavior of people attending ART clinics in Dar es Salaam with  
43 members of their social networks. In: 20th International AIDS Conference; 2014..  
44  
45  
46 [28] Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC medical research*  
47 *methodology*. 2006;6(1):54.  
48  
49  
50  
51  
52

- 1  
2  
3  
4  
5 [29] Zou J, Yamanaka Y, John M, Watt M, Ostermann J, Thielman N. Religion and HIV in Tanzania:  
6 influence of religious beliefs on HIV stigma, disclosure, and treatment attitudes. *BMC public health*.  
7 2009;9(1):75.  
8  
9  
10 [30] McCarthy B. Appreciative Inquiry: an alternative to behaviour management. *Dementia*.  
11 2017;16(2):249–253.  
12  
13  
14 [31] Siril HN, Kaaya SF, Fawzi MKS, Mtisi E, Somba M, Kilewo J, et al. CLINICAL outcomes and loss to  
15 follow-up among people living with HIV participating in the NAMWEZA intervention in Dar es  
16 Salaam, Tanzania: a prospective cohort study. *AIDS research and therapy*. 2017;14(1):18.  
17  
18  
19 [32] Carey MP, Schroder KE. Development and psychometric evaluation of the brief HIV Knowledge  
20 Questionnaire. *AIDS education and prevention*. 2002;14(2):172–182.  
21  
22  
23 [33] Haile BJ, Chambers JW, Garrison JL. Correlates of HIV knowledge and testing: Results of a 2003  
24 South African survey. *Journal of Black Studies*. 2007;38(2):194–208.  
25  
26  
27 [34] Aarø LE, Breivik K, Klepp KI, Kaaya S, Onya HE, Wubs A, et al. An HIV/AIDS knowledge scale for  
28 adolescents: item response theory analyses based on data from a study in South Africa and Tanzania.  
29 *Health education research*. 2011;26(2):212–224.  
30  
31  
32 [35] Easterly W. How the millennium development goals are unfair to Africa. *World development*.  
33 2009;37(1):26–35.  
34  
35  
36 [36] Barbarin OA, Khomo N. Indicators of economic status and social capital in South African townships:  
37 What do they reveal about the material and social conditions in families of poor children? *Childhood*.  
38 1997;4(2):193–222.  
39  
40  
41 [37] Lesko CR, Edwards JK, Cole SR, Moore RD, Lau B. When to censor? *American journal of epidemi-*  
42 *ology*. 2017;187(3):623–632.  
43  
44  
45 [38] VanderWeele T. *Explanation in causal inference: methods for mediation and interaction*. Oxford  
46 University Press; 2015.  
47  
48  
49  
50  
51  
52

- 1  
2  
3  
4  
5 [39] Forastiere L, Mealli F, Vanderweele TJ. Identification and Estimation of Causal Mechanisms in  
6 Clustered Encouragement Designs: Disentangling Bed Nets Using Bayesian Principal Stratification  
7 Identification and Estimation of Causal Mechanisms in Clustered Encouragement Designs : Disentan-  
8 gling Be. *Journal of the American Statistical Association*. 2016;111(514):510–525. Available from:  
9 <http://dx.doi.org/10.1080/01621459.2015.1125788>.  
10  
11  
12  
13  
14 [40] Shalizi CR, Thomas AC. Homophily and contagion are generically confounded in observational social  
15 network studies. *SMR*. 2012;40(2):211–39.  
16  
17  
18 [41] De La Haye K, Robins G, Mohr P, Wilson C. Homophily and contagion as explanations for weight  
19 similarities among adolescent friends. *Journal of Adolescent Health*. 2011;49(4):421–427.  
20  
21  
22 [42] Vanderweele TJ, Hong G, Jones SM, Brown JL. Mediation and spillover effects in group-randomized  
23 trials: a case study of the 4Rs educational intervention. *Journal of the American Statistical Association*.  
24 2013 jun;108(502):469–482. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23997375>.  
25  
26  
27  
28 [43] Latkin CA, Davey-Rothwell MA, Knowlton AR, Alexander KA, Williams CT, Boodram B. Social  
29 network approaches to recruitment, HIV prevention, medical care, and medication adherence. *Journal*  
30 *of Acquired Immune Deiciency Syndromes*. 2013;63:1–9.  
31  
32  
33 [44] Galea S, Tracy M. Participation rates in epidemiologic studies. *Annals of epidemiology*.  
34 2007;17(9):643–653.  
35  
36  
37 [45] Langford CPH, Bowsher J, Maloney JP, Lillis PP. Social support: a conceptual analysis. *Journal of*  
38 *advanced nursing*. 1997;25(1):95–100.  
39  
40  
41 [46] Schneider JA, Zhou AN, Laumann EO. A new HIV prevention network approach: sociometric peer  
42 change agent selection. *Social science & medicine*. 2015;125:192–202.  
43  
44  
45 [47] Leung MP. Treatment and spillover effects under network interference. *Review of Economics and*  
46 *Statistics*. 2016;p. 1–42.  
47  
48  
49  
50  
51  
52

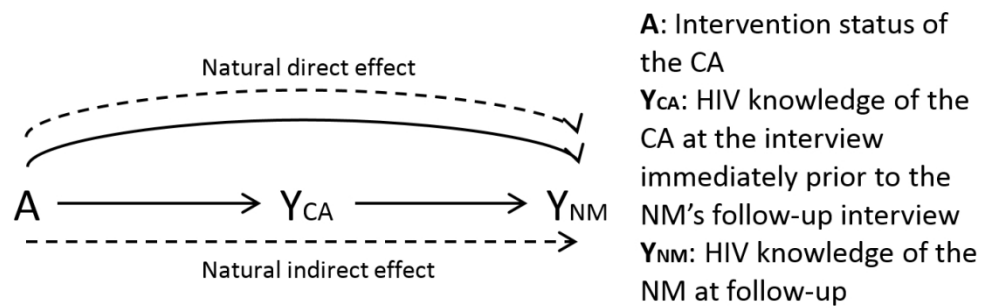
## 10. Figure Captions

Fig 1. Schematic of natural direct effect (NDE) and natural indirect effect (NIE). The NDE indicates the increase in NMs' HIV knowledge happens without a concomitant increase in their CA's HIV knowledge. The NIE, on the other hand, indicates that the increase in NMs' HIV knowledge is mediated by their CA's HIV knowledge increasing.

For peer review only



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



## STROBE (Strengthening The Reporting of OBServational Studies in Epidemiology) Checklist

A checklist of items that should be included in reports of observational studies. You must report the page number in your manuscript where you consider each of the items listed in this checklist. If you have not included this information, either revise your manuscript accordingly before submitting or note N/A.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).

Section and Item	Item No.	Recommendation	Reported on Page No.
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	
<b>Introduction</b>			
Background/Rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
<b>Methods</b>			
Study Design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed  <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	

Section and Item	Item No.	Recommendation	Reported on Page No.
Data Sources/ Measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study Size	10	Explain how the study size was arrived at	
Quantitative Variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	
Statistical Methods	12	(a) Describe all statistical methods, including those used to control for confounding	
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive Data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome Data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	

Section and Item	Item No.	Recommendation	Reported on Page No.
Main Results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other Analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key Results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
<b>Other Information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Once you have completed this checklist, please save a copy and upload it as part of your submission. DO NOT include this checklist as part of the main manuscript document. It must be uploaded as a separate file.**

# BMJ Open

## Evaluating spillover of HIV knowledge from study participants to their network members in a stepped-wedge behavioral intervention in Tanzania

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-033759.R2
Article Type:	Original research
Date Submitted by the Author:	26-Feb-2020
Complete List of Authors:	Lienert, Jeffrey; University of Pennsylvania Perelman School of Medicine, Center for Health Care Innovation; Philadelphia VAMC, CHERP Fawzi, Mary C; Harvard Medical School, Global Health and Social Medicine McAdam, Keith; London School of Hygiene and Tropical Medicine, Epidemiology Kaaya, Sylvia; Muhimbili University of Health and Allied Sciences, School of Medicine Liu, Yuanyuan Todd, Jim; London School of Hygiene and Tropical Medicine, Population Health Andrew, Irene; Management and Development for Health, Onnela, Jukka Pekka; Harvard T.H. Chan School of Public Health, Biostatistics
<b>Primary Subject Heading</b>:	HIV/AIDS
Secondary Subject Heading:	Global health, Public health
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, STATISTICS & RESEARCH METHODS

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3  
4  
5 Evaluating spillover of HIV knowledge from study participants to their  
6 network members in a stepped-wedge behavioral intervention in Tanzania  
7  
8

9 Jeffrey Lienert<sup>a,b,j</sup>, Mary C. Smith Fawzi<sup>c</sup>, Keith McAdam<sup>d</sup>, Sylvia Kaaya<sup>e</sup>, Yuanyuan Liu<sup>i</sup>, Jim Todd<sup>g</sup>,  
10 Irene Andrew Lema<sup>f</sup>, Jukka-Pekka Onnela<sup>h</sup>

11  
12 <sup>a</sup>Perelman School of Medicine, University of Pennsylvania <sup>b</sup>Center for Health  
13 Equity Research and Promotion, Philadelphia VA <sup>c</sup>Department of Global Health  
14 and Social Medicine, Harvard Medical School <sup>d</sup>Department of Clinical Tropical  
15 Medicine, London School of Hygiene and Tropical Medicine <sup>e</sup>Department of Psychiatry and  
16 Mental Health, Muhimbili University of Health and Allied Sciences <sup>f</sup>Management and  
17 Development for Health <sup>g</sup>Department of Population Health, London School of  
18 Hygiene and Tropical Medicine <sup>h</sup>Department of Biostatistics, Harvard T. H. Chan School of  
19 Public Health <sup>i</sup>University of Texas School of Public Health  
20 <sup>j</sup>Corresponding Author: Jeffrey Lienert, jeffrey.lienert@penncmedicine.upenn.edu  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

52 Preprint submitted to BMJ Open

25 February, 2020

## Abstract

**Objectives:** We aim to describe the social network members of participants of a behavioral intervention, and examine how the effects of the intervention may spillover among network members.

**Design:** Secondary analysis of a step-wedge randomized controlled trial.

**Setting:** Change Agents (CAs) were recruited from waiting rooms of HIV treatment facilities in Dar es Salaam, Tanzania, and their Network Members (NMs) were recruited directly by CAs.

**Participants:** We enrolled 662 CAs in an HIV behavioral intervention. They, along with 710 of their NMs, completed baseline and follow-up interviews from 2011 to 2013.

**Primary and Secondary Outcomes:** The primary outcome of this study was change in NMs' HIV knowledge, and the secondary outcome was whether the NM was lost to follow-up.

**Results:** At baseline, many characteristics were different between NMs and CAs. We found a number of NM characteristics significantly associated with follow-up of NMs, particularly female gender (OR=1.64, 95% CI: 1.02,2.63) and HIV knowledge (OR=20.0, 95% CI: 3.70,125); only one CA variable was significantly associated with NM follow-up: having a private source of water (OR=2.17, 95% CI: 1.33,3.57). The 14.2% increase in NMs' HIV knowledge was largely due to CAs feeling empowered to pass on prior knowledge, rather than transmitting new knowledge to their NMs.

**Conclusions:** Characteristics of social network members of persons living with HIV PLH may play a role in study retention. Additionally, the HIV knowledge of these NMs increased largely as a function of CA participation in the intervention, suggesting that intervening among highly-connected individuals may maximize benefits to the potential population for whom spillover can occur.

*Keywords:*

mediation, natural direct effect, natural indirect effect, follow-up, retention, hidden population



## 1. Strengths and Limitations of This Study

- The recruitment method, focused on social network members of persons living with HIV, allowed us to access an at-risk population typically not easily-accessible via other means, such as direct approach by the study team rather than someone known to the participant.
- The study design of the trial in which this study is nested was an ideal situation in which to employ a mediation analysis that informs our understanding of how this and similar intervention effects may spillover in a population.
- The greater-than-ideal dropout rate of NMs was both a strength in that it allowed us to examine factors associated with dropout, but also a limitation, in that the potential of differential dropout by unmeasured factors may have biased some of our results.

## 2. Introduction

Despite recent decreases in mortality, HIV/AIDS is currently the ninth leading cause of years of life lost (YLL) globally [1]. Research into eliminating HIV has focused on two fronts: biological and behavioral, which combined have decreased the years of life lost due to HIV by an estimated 51% [1]. Behavioral interventions seek to curb transmission of HIV through reduction of risk behavior including safer sexual practices and reducing injection drug use [2, 3]. Behavioral change also impacts the effectiveness of the biological methods, given that lower adherence is known to often compromise their effectiveness [4]. However, neither biological nor behavioral methods are fully effective without the knowledge of these treatments and preventative behaviors. It is therefore important to increase HIV-related knowledge as a necessary, but not sufficient, step towards achieving the 90-90-90 treatment goals (90% diagnosis, 90% antiretroviral therapy, and 90% viral suppression among the treated) and ultimately eliminate HIV [5].

The importance of adequate HIV knowledge has been recognized for a period of time, as it is often a prerequisite for behavioral change [6]. It should be noted that increased knowledge does not necessarily translate into behavior change, as other factors such as motivation and skills also play a role [6]. Despite this, researchers conducted trials among Persons Living with HIV (PLH) which led to increased HIV knowledge [7, 8]. These interventions were group discussions and teaching led by nurses, and motivational interviewing, respectively. Researchers also found that the intervention reduced HIV-risk behaviors

1  
2  
3 concomitantly with an increase in HIV knowledge [8].  
4

5 Knowledge gained by participants in these trials can also be freely shared with members of their social  
6 network, in what is known as a spillover effect [9]. Specifically, a spillover effect (also known as an in-  
7 direct or disseminated effect) is one person's exposure affecting another's outcome [10]. In the context of a  
8 social network spillover intervention, it means a change to someone's behavior who did not receive the  
9 intervention because they were socially connected to someone who did receive the intervention. This is dis-  
10 tinct from what is sometimes called behavioral spillover, where change's in a person's behavior affects other  
11 behaviors of that same person [11]. For injection drug users, HIV prevention educational interventions were  
12 demonstrated to have spillover effects of HIV prevention education, and subsequent reduced rates of risky  
13 behaviors [12]. What remains unknown, however, is whether or not a spillover effect exists for HIV  
14 knowledge during and after an intervention in other populations, particularly sub-Saharan Africa. In other  
15 words, we aim to determine that if an intervention increases someone's HIV knowledge, how members of  
16 their social networks also increase their HIV knowledge.  
17  
18  
19  
20  
21  
22  
23  
24

25 New knowledge can come from a variety of sources, one of the most important of which is a person's  
26 social network [13, 14]. Social networks are of particular import because new knowledge can lead to  
27 cascades of behavior change, where people subsequently educate those in their social network, in what is  
28 known as social influence [15, 16, 17]. This has been examined in participant-driven interventions, where  
29 initially-recruited participants educate members of their social network one-on-one [18]. Characteristics  
30 such as knowing the HIV status of network members has been shown to be the most important predictor of  
31 engaging in prevention advocacy [19]. Work on diffusion through social networks, how a belief or behavior  
32 can be "contagious" within a network, has shown that spreading intervention effects beyond the initial study  
33 population can improve the cost-effectiveness of these interventions [20]. These findings imply that certain  
34 aspects of knowledge or behavior may spread more or less efficiently through networks comprising  
35 individuals with specific characteristics, which may need to be accounted for in network interventions. For  
36 instance, networks comprising many at-risk individuals who are HIV-negative may not be as receptive to  
37 change in behaviors as networks comprising a mix of those with and without HIV. Additionally, the  
38 networks of PLH are often difficult networks to ascertain [21], due to the continued stigma of HIV and AIDS  
39 in many settings [22]. Because of this, if a PLH or person at-risk of acquiring HIV does not want to  
40 participate in an investigator-initiated intervention, there is little recourse other than information transmitted  
41 via social networks, or targeted sampling techniques which are not always effective (e.g. Respondent Driven  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53

1  
2  
3 Sampling) [23]. This is particularly important in low- and middle-income countries, as a recent systematic  
4 review found only 54 studies researching spillover effects in these settings (out of approximately 750) [24].  
5  
6 Therefore, understanding exactly how information spreads from participants in an intervention to members  
7 of their social network, who may be largely inaccessible via other means, is important for reaching the  
8 greatest number of people about HIV prevention. Understanding what makes these persons different from  
9 those who receive the intervention themselves is important, as it may point to ways in which to increase  
10 enrollment of these populations.  
11  
12  
13  
14

15 Based on the above gaps in the literature, we conducted a study on network members of PLH enrolled  
16 in a behavior change intervention [25]. The trial recruited PLH to serve as Change Agents (CAs) and to  
17 reach out to their social network members (NMs) about knowledge of HIV and safer sexual practices [26].  
18 Our goals in this study were threefold: 1) to understand how the NMs differed from their CAs, 2) understand  
19 correlates of dropout for the NMs, and 3) to understand by what method HIV knowledge transferred to the  
20 NMs from the CAs. Understanding how the information and behaviors are shared within social networks  
21 will allow HIV researchers and others to take advantage of this knowledge and improve upon prevention  
22 interventions in the future.  
23  
24  
25  
26  
27  
28  
29

### 30 **3. Methods**

#### 31 *3.1. Study Population*

32 We analyzed social network data from the Agents of Change trial [27], which was a stepped-wedge  
33 randomized controlled trial [28] that enrolled PLH to become Change Agents (CAs) by informing members  
34 of their social network (NMs) about knowledge of HIV and safer sexual practices. Here, we define CAs  
35 based on the *potential* for PLH to become so - by self-selecting into the study, PLH identify themselves as  
36 potential CAs, which we aim to foster through the intervention.  
37  
38  
39  
40  
41

42 CAs were recruited from the waiting rooms of HIV care and treatment centers in Dar es Salaam, Tanza-  
43 nia, and we received written consent from each CA. Participants were given a baseline questionnaire and  
44 were randomized to one of three waves in which to receive the intervention. At baseline, participants were  
45 also asked to recruit up to three members of their social networks who they felt were at particularly high risk  
46 of contracting or spreading HIV. We obtained written consent from these nominated network members (NMs).  
47 NMs could be either HIV positive or negative, and they were given a baseline survey. The NM was only  
48 aware the CA was a participant in the intervention if the CA shared this information with them, which  
49  
50  
51  
52  
53

1  
2  
3 many did not due to HIV-related stigma [29]. Each CA therefore formed a CA-NM dyad with each NM they  
4 recruited, and if they recruited more than one NM, formed a set of CA-NM dyads with a common CA.  
5

6 For CAs, the intervention comprised 10 weekly structured sessions aimed at empowering PLH to be-  
7 come HIV prevention change agents in their communities. The sessions aimed to increase psychosocial and  
8 communication skills through an Appreciative Inquiry approach [30]. Within one month of each wave of the  
9 intervention, CAs were given follow-up surveys. Across all waves, the interventions lasted from November  
10 2010 to January 2014, and the final interviews were conducted in March 2014.  
11  
12

13 The NMs did not receive any intervention at any point during the study. Rather, their intervention status  
14 flowed from the intervention status of their CAs. Due to operational difficulties, each NM was surveyed two  
15 times during the study: baseline and after the first wave, rather than baseline and one after each wave. We  
16 use this interview of NMs as the division between “exposed” and “unexposed”. At the time of an NM’s  
17 follow-up interview, their respective CA may or may not have undergone the intervention. Therefore, the  
18 NMs were divided into “exposed” (N=381) and “unexposed” (N=329) groups based on whether their  
19 respective CA had completed their intervention at the time of the NM’s follow-up interview. In this way,  
20 we were able to assess the spillover effect of the intervention net of temporal or geographical trends.  
21  
22

23 Inclusion criteria for the follow-up analyses were for CAs who completed at least one follow-up inter-  
24 view and had at least one NM who completed at least one follow-up questionnaire. In this way, outcomes  
25 could be computed for the CA and NM of each dyad. NMs approached by a CA, but who did not participate  
26 in the study were not recorded since it was not feasible to obtain this information from study participants  
27 themselves. As we lost some CAs and NMs to follow up, we completed our analyses without their data,  
28 assuming it to be Missing Completely at Random (MCAR). In sum, our sample comprises 662 CAs and 710  
29 NMs, meaning each CA recruited 1.07 NMs on average out of a possible 3, and 44 CAs nominated at least  
30 two NMs.  
31  
32

33 During this study, there was little loss-to-follow-up among the CAs (< 10%), but much higher among  
34 the NMs (36.8%) [31].  
35  
36

### 37 3.2. *HIV knowledge*

38 To assess HIV knowledge of CAs and NMs, we used the brief HIV knowledge questionnaire [32]. This  
39 scale comprises 18 questions, which focus on an individual’s knowledge of how HIV can spread and other  
40 characteristics of the virus and AIDS. This knowledge is crucial for an individual’s subsequent safe-sex  
41  
42

1  
2  
3 practices to reduce the risk of transmission. The original population comprised three different groups: two  
4 groups of low-income women, and one of women and men receiving psychiatric treatment. In these popula-  
5 tions, the measure had a Chronbach's alpha of 0.78. This instrument has been used previously in sub-  
6 Saharan Africa and has demonstrated reasonable reliability of 0.75 Cronbach's alpha in South Africa among  
7 a convenience sample of 429 members of the African Methodist Episcopal church [33]. It has also been  
8 translated to Swahili, with only minor differential item functioning [34]. This indicates that the measure  
9 performs adequately in other, similar populations, relevant to this work. Here, the Cronbach's alpha was  
10 0.71. Because the average baseline score of HIV knowledge was 0.80, with a preponderance of scores of  
11 1.00, the main indicator of knowledge was not normally distributed, and therefore a continuous predictor  
12 was not ideal [35,36]. We therefore summarize this measure as "Complete HIV Knowledge", a dichotomous  
13 variable for whether the participant correctly answered all questions.  
14  
15  
16  
17  
18  
19  
20  
21

### 22 3.3. *Demographic and contextual variables*

23  
24  
25 In addition to HIV knowledge, we are interested in a set of demographic and contextual variables that  
26 may help explain some of the trends observed. In terms of demographic variables, we include age, sex,  
27 employment status, marital status, and self-identified HIV status. We also include education, which we di-  
28 chotomize at 7 years or more - a level commensurate with elementary education. We selected this cutoff  
29 because it coincides with the millennium development goal (MDG) of increasing primary education com-  
30 pletion [37].  
31  
32  
33  
34

35 Contextually, we included two additional variables: having a private source of water, and the number  
36 of persons sleeping in the participant's home. The first was also based on a MDG, and indicates participants  
37 with access to safe drinking water [37]. This is a proxy for the economic security of the participant. The  
38 number of persons sleeping in the participant's home is also a non-monetary indicator of their material and  
39 social resources [38]. These variables combined give a more thorough picture of the participant's economic  
40 status than employment alone.  
41  
42  
43  
44

### 45 3.4. *Statistical Analysis*

46  
47 The first analysis was to assess homophily, defined as the extent to which CAs and NMs were similar to  
48 one another in terms of a number of sociodemographic characteristics and HIV-related risk factors. Because  
49 CAs and NMs self-selected into the study and were not randomized to CA/NM status, we do not *a priori*  
50  
51  
52

1  
2  
3 expect them to be completely similar. In addition, multiple NMs could share a CA and would therefore not  
4 be independent due to the shared variation and latent characteristics of having the same CA. We therefore  
5 assessed statistical significance of homophily on the set of CA-NM dyads using a permutation test. For  
6 continuous variables, the difference between the CA and NM was calculated, and for categorical variables,  
7 whether the CA and NM were concordant or discordant was recorded. We then randomly permuted CA-NM  
8 ties (keeping number of ties per CA constant), and then recalculated these statistics 1,000 times. We then  
9 examined the percentile of the observed difference relative to the permuted differences [39]. Analyses were  
10 run using R v3.1.1.

11  
12 To accomplish our second aim of understanding what was associated with NMs completing their follow-  
13 up interview, we fit a log-binomial regression to determine predictors of follow-up. In this regression, we  
14 use all the variables listed above, as well as whether the CA remained in the study for its full duration. In  
15 order to examine the association between these same variables and the time of follow-up, we also employed  
16 a Cox regression. In the Cox regression, start of follow-up was defined as the time at which the NM  
17 completed their baseline interview. The outcome here was whether the NM completed a follow-up interview.  
18 NMs who were lost to follow-up were censored at the time of their latest interview [40]. For this analysis  
19 we report Hazard Ratios (HRs), which have important limitations: selection bias and sensitivity to study  
20 period [41]. Either of these could affect this analysis, hence our use of logistic regression as a primary  
21 analysis. However, they remain useful as a sensitivity analysis.

22  
23 Finally, since the trial showed beneficial effects on the HIV knowledge of the NMs following the inter-  
24 vention [27], we aimed to elucidate exactly what caused the HIV knowledge of NMs to increase - either  
25 CAs gaining knowledge through the intervention and passing it on, or the CAs being empowered by the  
26 intervention to pass on existing knowledge. As the wedge in which the CA received the *NAMWEZA*  
27 intervention was randomized, we treat the time at which an NM was potentially exposed to *NAMWEZA*  
28 through their CA as similarly randomized. This randomization scheme allowed us to explore the spillover  
29 effect of CAs' HIV knowledge onto their respective NMs via a mediation analysis.

30  
31 As shown by VanderWeele et al (2015), social network spillover effects in the case of dyadic  
32 relationships can be broken down into concepts from mediation analysis: direct and indirect effects (Figure  
33 1) [42]. This method has since been used for novel evaluations of spillover effects [43]. Although previous  
34 studies showed that this same type of analysis cannot be done on full network data, the data in this study  
35 consisted of only dyads, the CAs and their NMs, so in this case the method is appropriate [44, 45].

1  
2  
3 One assumption of this analysis is that the dyads are independent, which is violated here; if a CA  
4 recruited more than one NM, the multiple CA-NM dyads involving the same CA would not be independent.  
5 To address this, we performed the analysis after randomly removing NMs until each CA had only a single  
6 NM. This resulted in removing 48 NMs, just 6.7% of the population. We found that the point estimates were  
7 nearly identical, but that the confidence intervals were slightly larger due to the reduced sample size. No  
8 coefficients changed from significant to non-significant in this analysis (data not shown).  
9

10  
11 A second, related assumption is that of *partial interference*, that the effects in one cluster does not affect  
12 another cluster - here, one CA-NM dyad affecting another [46]. This could occur if two NMs of different  
13 CAs happen to know one-another outside of the study, one has a CA who was randomized to an earlier  
14 wedge, and shares what they know of it with the otherwise-unexposed NM. However, due to the size of Dar  
15 es Salaam, and the number of HIV treatment clinics in which recruitment occurred, we expect few CAs or  
16 NMs to know one another outside of the study, limiting the potential for partial interference.  
17  
18

19  
20 The method parses social influence into direct and indirect effects. The natural indirect effect (NIE) is  
21 the effect by which the intervention changes the CA's HIV knowledge, which then changes the NM's HIV  
22 knowledge. The NIE is similar to the effect observed in many participant-driven interventions: an initial  
23 participant receives the intervention, increasing their knowledge, and they subsequently pass their increased  
24 knowledge to members of their social network [18]. The natural direct effect (NDE) is the effect of receiving  
25 the intervention has on an NMs outcome, irrespective of the CA's outcome (in this case HIV knowledge).  
26 For instance, this could occur if CAs begin with good knowledge of HIV, and the intervention empowers  
27 them to convey knowledge they already had to their NMs. Although the intervention does not increase their  
28 HIV knowledge, it is still useful to the CAs, as it empowers them to become CAs in the first place. In order  
29 to estimate these effects, the published SAS macro developed by VanderWeele and colleagues was used,  
30 and analyses were run using SAS v9.2 (SAS Institute, Cary, NC) [47,48]. The data are not publicly available  
31 due to the sensitive nature of HIV infection status.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

### 44 3.5. Patient and Public Involvement

45  
46 Patients were not involved in the design of this study. As part of the recruitment process, patients were  
47 instructed to recruit members of their social networks who they felt were at high risk of HIV infection - in  
48 this way participants could guide the recruitment of NMs to those most at-risk. We currently have no plans  
49 to disseminate the results of this study to participants.  
50  
51

#### 4. Results

The *NAMWEZA* study recruited 662 Change agents (CA), of whom 453 (68.4%) completed at least one follow-up questionnaire. These 662 CAs in turn recruited 710 network members (NM) who took the baseline questionnaire. Of the 710 NMs, 449 (63.2%) also completed a follow-up questionnaire (Table 1). At baseline, CAs were on average older than their NMs. More of the NMs were employed than their CAs (69.3% vs. 55.0%,  $p<0.001$ ), but were less likely to have at least 7 years of education (52.0% vs. 52.3%,  $p<0.001$ ). Only 12.3% of NMs were HIV-positive, compared to *all* CAs ( $p<0.001$ ). Complete data was obtained at baseline for all CAs and NMs.

Characteristic	Number of NMs (%) or Mean (SD) (N=710)	Number of CAs (%) or Mean (SD) (N=662)	P-value
Age	33.0 (11.1)	38.9 (9.7)	<0.001
Female	380 (53.7%)	349 (53.9%)	0.89
Employed	490 (69.3%)	356 (55.0%)	<0.001
At least 7 years education	369 (52.0%)	584 (82.3%)	<0.001
Complete HIV knowledge	638 (89.9%)	598 (90.4%)	0.65
Persons sleeping in home	5.00 (5.77)	3.87 (3.72)	<0.001
Married	373 (52.7%)	338 (51.1%)	0.56
HIV Positive	87 (12.3%)	662 (100%)	<0.001
Private source of water	309 (43.7%)	263 (39.7%)	0.19

Table 1: Demographic characteristics at baseline, with the results of a permutation test for homophily. Specifically, for each CA- NM dyad, either a difference (for continuous variables) or concordance (for dichotomous variables) is calculated. For example, if a CA was 39 years old, and their NM was 25 years old, the difference would be 14 years old. If a CA was male and their NM was Female, the pair would be discordant for sex. CA-NM pairs were then randomly reshuffled, the edge-wise characteristics recalculated, and the observed difference compared to the distribution of randomized differences.

Logistic regression showed that characteristics of both the CAs and the NMs significantly predicted loss to follow-up (Table 2). The NM being female (RR=1.44, 95% CI: 1.05,1.97), having complete HIV knowledge (RR=10, 95% CI: 2.33,42), being employed (RR=1.43, 95% CI: 1.08,1.89), and being married (RR=1.55, 95% CI: 1.03,2.33) were all significant predictors of increased odds of completing a follow-up interview. Each additional person sleeping in the home of the NM per room used for sleeping reduced the odds of follow-up (RR=0.85, 95% CI: 0.74,0.98) as did the NM living with HIV (RR=0.40, 95% CI: 0.17,0.96). CA having a private water source was significantly associated with increased odds of the NM being followed-up (RR=2.07, 95% CI: 1.25,3.42), even after controlling for the NM having a private source



of water. This was the only CA-specific variable that significantly predicted an NM's follow-up.

The Cox proportional hazard model showed similar results to the logistic regression, but with fewer significant results. Only the NM being married (HR=1.28, 95% CI: 1.01,1.64) and the NM living with HIV (HR=0.71, 95% CI: 0.51,0.99) significantly predicted the time-to-follow-up of the NMs. Additionally, all of the hazard ratios were closer to the null relative to the corresponding odds ratios.

We assess the mechanism of spillover effects on HIV knowledge, the main effect of the intervention on the NMs. The analysis showed that NMs who were exposed to intervention via their CA had an increase in HIV knowledge of 12.9% on average (95% CI: 0.06, 0.20) from a baseline percentage of 78% [26]. This was broken down into a natural indirect effect (NIE) of 0.6% (95% CI: -0.06%,2.0%), which is the effect the intervention had via the CA's HIV knowledge, and a natural direct effect (NDE) of 12.3% (95% CI:

Characteristic (N=459)	Adjusted RR (95% CI)	Adjusted HR (95% CI)
Characteristics of NMs		
Gender (Female)	1.44* (1.05,1.97)	1.18 (0.94,1.50)
Difference in age of NM and CA (per year)	1.01 (0.99,1.03)	1.01 (0.995,1.02)
Age of NM	0.98 (0.95,1.02)	0.99 (0.97,1.04)
Complete HIV knowledge (vs. none)	10* (2.33,42)	2.20 (0.97,5.01)
Employed	1.43* (1.08,1.89)	1.15 (0.89,1.50)
Each additional person sleeping in home per room used for sleeping	0.85* (0.74,0.98)	0.92 (0.83,1.01)
Married	1.55* (1.03,2.33)	1.28* (1.01,1.64)
Living with HIV	0.40* (0.17,0.96)	0.71* (0.51,0.99)
Having a private source of water	0.89 (0.59,1.34)	0.97 (0.77,1.22)
Characteristics of CAs		
Gender (Female)	1.27 (0.76,2.08)	1.07 (0.84,1.37)
Complete HIV knowledge (vs. none)	0.36 (0.07,2.04)	0.64 (0.29,1.43)
Having a private source of water	2.07* (1.25,3.42)	0.97 (0.77,1.22)
Being employed	1.54 (0.95,2.50)	1.14 (0.90,1.43)
Being married	1.25 (0.77,2.04)	1.11 (0.88,1.43)
CA Lost to follow-up	1.06 (0.84,1.33)	1.02 (0.72,1.43)

Table 2: Results of multivariate log-binomial regression and Cox-proportional hazard models on the dichotomous outcomes of whether the NM completed a follow-up questionnaire, and the continuous outcome of time-to-completion of follow-up questionnaire, respectively. “\*” indicates significant at the  $p < 0.05$  level.

6.1%,19.3%), which is the effect the CA's participation had on the NM's HIV knowledge, irrespective of the CA's HIV knowledge. In other words, of the 12.9% increase in NM's having complete HIV knowledge, 12.3% occurred without a concomitant increase in their CA's HIV knowledge. In other words, their HIV knowledge increase was not mediated by the increase in HIV knowledge of their CA. This did not change when we used only one NM per CA.

## 5. Discussion

Based on results from the NAMWEZA trial in Tanzania, we have shown a number of novel findings regarding the network members of HIV intervention participants, correlates of their retention in the study, and the evidence for spillover of HIV knowledge from CAs to NMs. These findings can inform the design of interventions in the future to maximally enroll and retain participants and ensure that intervention information is transmitted from the study participants to members of their social networks.

We found several significant differences between CAs and NMs they recruited. On average, CAs were older, less likely to be employed, and more educated. All of this suggests that by having CAs recruit from their social network, we were able to recruit a set of social network members different in multiple ways from those found in the waiting rooms in HIV treatment clinics (unsurprising, given that many NMs were not HIV-positive) [9]. This may in part be due to the fact that when recruiting NMs, CAs were instructed to recruit those who may be at particular risk of contracting or spreading HIV. Therefore, CAs likely did not pick random members of their social network, but those who were at high risk. This finding corroborates the claim by Latkin et al. (2013) that sampling via enumeration of network members by initial participants recruits a more diverse sample, as our sample of NMs was not composed of only those who were HIV-positive, but also many who were HIV-negative [49]. Additionally, the lower likelihood of employment of the CAs indicates they may have had more time for participating in the intervention [50]. This means that future studies may need to tailor their interventions to work with the schedules of employed persons to increase participation.

Although the study design potentially accessed a separate slice of the population than other methods vis-a'-vis the population of those at risk for contracting HIV, it may have been at least partially responsible for the large loss of follow-up of the NMs. It is important to note that more NM characteristics than CA characteristics were significantly predictive of follow-up in both the logistic and Cox proportional hazards models. The one CA characteristic which did predict NMs completing a follow-up interview, having a private source of water, was a proxy for CA's socioeconomic status. This may have been because CAs with greater resources may have had more time available to pass information to their NMs, retaining the NM's interest longer [51]. We also note that less HIV knowledge and living with HIV predicted reduced likelihood of completing a follow-up questionnaire among NMs, which means that those who might have benefited most from spillover of the intervention were more likely to discontinue their involvement. This does not

1  
2  
3 mean that they did not receive any spillover, only that it was not recorded. Our estimate of the magnitude of  
4 the spillover may therefore be biased towards the null. This is also problematic more generally for  
5 interventions of this nature as the very people the intervention aims to benefit may not stay with the program.  
6  
7

8 The HIV knowledge gain experienced by the NMs was largely due to the NDE; i.e. knowledge spilled-  
9 over as a result of the CAs participating in NAMWEZA, not because the CAs' HIV knowledge increased,  
10 and they passed this new knowledge to their NM(s). This is most likely because CAs had a high average  
11 HIV knowledge score at baseline (80%), so there was less room for improvements in knowledge. Following  
12 the intervention, CAs' likelihood of complete HIV knowledge did not significantly increase. Therefore, what  
13 prompted the increased knowledge of the NMs was the CA becoming empowered through the intervention  
14 to pass on their existing knowledge to their NM [52].  
15  
16  
17  
18  
19

20 This finding is important for future interventions: spillover effects of this intervention will likely carry  
21 over only to those directly-connected to the CAs (opposed to spreading indefinitely in a snowball effect), as  
22 the most important thing in transmitting HIV knowledge is receiving the intervention, which only CAs  
23 receive. Because NMs do not actually receive the intervention, it is unlikely that they would feel empowered  
24 to become CAs themselves (particularly because they will not receive the *NAMWEZA* sessions), thus limiting  
25 the spread of the intervention. For CAs to increase the HIV knowledge of their NMs, they would only need  
26 to become empowered to share their information. However, for the NMs to increase the HIV knowledge of  
27 their own NMs, they would need to both increase their HIV knowledge, and become empowered to share it.  
28 This is less likely to happen than just becoming empowered, and so it is unlikely, though not impossible, that  
29 this effect would continue to spread in the population. This may give insight into how to design interventions  
30 in the future; if one wants to maximize the number of people who benefit, choosing CAs who form many  
31 bridging ties in the community would maximize the potential number of links by which spillover can occur  
32 [53]. Alternatively, interventions can be designed to be self-propagating; if CAs are empowered to deliver  
33 the intervention to others, changing their own NMs into future CAs, those new NMs-turned-CAs could then  
34 deliver the intervention to a second set of NMs, again empowering them to become CAs.  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

45 Our study is not without limitations. First, there was considerable loss to follow-up of the NMs. Al-  
46 though this actually informed our analysis of the correlates of loss to follow-up, it meant that our final  
47 analysis may have biased our results. Even though the exposure was randomized, the loss to follow-up can  
48 result in selection bias if the NMs who left the study were systematically different from those who remained.  
49 As we show, the NMs who dropped out were those who would have benefited the most from the intervention  
50  
51  
52

1  
2  
3 because they were more vulnerable and at higher risk than those who did not drop out. Because they are  
4 likely to have greater benefits from any spillover effects, we would expect this non-random loss to follow-up  
5 to result in an underestimate of the impact of the intervention in the network members. Adjusting for  
6 censoring weights may ameliorate this issue [54]. Second, our data did not perfectly fit the requirements of  
7 the causal mediation analysis: they were not entirely independent, since multiple NM-CA dyads shared a  
8 CA. However, when we randomly removed dyads until there were no repeated CAs, the results were  
9 qualitatively very similar, indicating that lack of independence did not unduly affect our results. Third,  
10 although we were able to tease apart the direct and indirect effects, we are unable to determine the mechanism  
11 of the natural direct effect; the data do not allow us to specify whether NMs increased knowledge through  
12 speaking to their knowledgeable CA, through researching HIV on their own, or some other mechanism.  
13 Future work will have to be done to examine these different pathways.  
14  
15  
16  
17  
18  
19  
20  
21  
22

## 23 **6. Conclusions**

24  
25 These results have implications for the potential scale-up of the *NAMWEZA* intervention, as well as  
26 future studies and interventions that focus on behavioral interventions in social networks. First, our findings  
27 of minimal similarity between CAs and their NMs indicate that this recruitment method allows us to enroll  
28 participants from portions of the population that are not represented by the CAs alone. Coupling this with a  
29 deeper understanding of the mental heuristics CAs used to select NMs (e.g. did CAs mentally search their  
30 close or peripheral network for those at-risk of HIV), may lead to different strategies for recruitment and  
31 retention, leading to stronger effects of behavioral interventions. The mediation analysis presents a  
32 compelling picture of how best to ensure the benefits of interventions reach as many people beyond the  
33 study participants as possible. Participation of CAs in the intervention resulted in positive effects on their  
34 immediate network members' HIV knowledge regardless of how the CAs responded to the intervention.  
35 While improvement in HIV knowledge may not necessarily translate to increased safe sex behavior, it can  
36 be seen as a rate-limiting step towards reducing HIV transmission since adequate knowledge is usually  
37 necessary for reducing risk behavior. Future work should examine the exact mechanisms of spillover  
38 discussed here, as that is an important clarification that could benefit future studies. Specifically, a similar  
39 setup to the study here combined with semi-structured interviews with CAs and NMs about their interactions  
40 with one another would help elucidate exactly how NMs increased their HIV knowledge (this population  
41 would likely need to comprise CAs who have disclosed their HIV status in order to prevent accidental  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53

disclosure). Future work should also examine whether increases in HIV knowledge translate to changes in behaviors which may increase one's risk of contracting HIV, particularly in Sub-Saharan Africa. The results presented herein may inform approaches for increasing participation and potentially conferring greater benefits related to spillover effects in future HIV behavioral interventions.

## 7. Acknowledgements

The authors would like to thank the study participants and staff for their time, hard work, and dedication. We would also like to thank Guy Harling for their comments on the manuscript.

## 8. Contributorship

JL, MCSF, and JPO designed the analysis plan and drafted the initial paper. JPL, MCSF, KM, SK, IAL and JT conducted the underlying randomized trial, and designed it to incorporate social networks data and spillover information; these same authors provided feedback on the proposed analysis plan. JL and YL conducted analyses for the paper. All authors helped revise the drafts of the paper and approved the final manuscript.

## 9. Compliance with ethical standards

**Funding:** The authors received funding from the CDC via grant TZ/UG.08.0147, the President's Emergency Plan for AIDS Relief - Tanzania, and the UTHealth Innovation for Cancer Prevention Research Training Program Pre-doctoral Fellowship (Cancer Prevention and Research Institute of Texas grant RP160015).

**Conflict of interest:** The authors declare that they have no conflict of interest.

**Ethical approval:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. IRB approval was obtained from Harvard Medical School, the Muhimbili University of Health and Allied Sciences, and the National Institute for Medical Research (NIMR) in Tanzania. Particular attention was paid to eliminate the risk of inadvertent disclosure of CA HIV status to NMs.

**Informed consent:** Informed consent was obtained from all individual participants included in the study.

**Disclaimer:** The content is solely the responsibility of the authors and does not necessarily represent the official views of the Cancer Prevention and Research Institute of Texas.

**Data sharing statement:** There are no additional data available. As the data contain enough information to potentially-uniquely identify specific participants, in the sensitive context of HIV, we have chosen not to make the data available.

## 10. References

- [1] Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*. 2017;390(10100):1151–1210.
- [2] Crepaz N, Tungol-Ashmon MV, Higa DH, Vosburgh W, Mullins MM, Barham T, et al. A systematic review of interventions for reducing HIV risk behaviors among people living with HIV in the United States, 1988–2012. *AIDS*. 2014 mar;28(5):633–56. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24983541>.
- [3] Arnold EA, Kegeles SM, Pollack LM, Neilands TB, Cornwell SM, Stewart WR, et al. A Randomized Controlled Trial to Reduce HIV-Related Risk in African American Men Who Have Sex with Men and Women: the Bruthas Project. *Prevention Science*. 2018;p. 1–11.
- [4] Garcia de Olalla P, Knobel H, Carmona A, Guelar A, Lopez-Colomes J, Cayla J. Impact of adherence and highly active retroviral therapy on survival in HIV-infected patients. *Journal of Acquired Immune Deficiency Syndromes*. 2002;30(1):105–10.
- [5] Joint United Nations Programme on HIV/AIDS and Joint United Nations Programme on HIV/Aids and others. 90-90-90: an ambitious treatment target to help end the AIDS epidemic. Geneva: UNAIDS. 2014;.
- [6] Fisher JD, Fisher WA. Changing AIDS-risk behavior. *Psychological bulletin*. 1992;111(3):455.
- [7] McNeill C, George N, Glover R. An Evaluation of Sisters Informing Healing Living Empowering: Increasing HIV Knowledge Among African American Adolescent Females Using an Evidence-Based HIV Prevention Intervention. *Journal of Doctoral Nursing Practice*. 2017;10(1):4–10.
- [8] Carey MP, Braaten LS, Maisto SA, Gleason JR, Forsyth AD, Durant LE, et al. Using information, motivational enhancement, and skills training to reduce the risk of HIV infection for low-income urban

- women: A second randomized clinical trial. *Health Psychology*. 2000;19(1):3.
- [9] Sinclair B, McConnell M, Green DP. Detecting spillover effects: Design and analysis of multilevel experiments. *American Journal of Political Science*. 2012;56(4):1055–1069.
- [10] Benjamin-Chung J, Arnold BF, Berger D, Luby SP, Miguel E, Colford Jr JM, et al. Spillover effects in epidemiology: parameters, study designs and methodological considerations. *International journal of epidemiology*. 2017;47(1):332–347.
- [11] Truelove HB, Carrico AR, Weber EU, Raimi KT, Vandenberg MP. Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental Change*. 2014;29:127–138.
- [12] Ghosh D, Krishnan A, Gibson B, Brown SE, Latkin CA, Altice FL. 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*. 2017;21(4):1183–1207.
- [13] Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *The New England journal of medicine*. 2007;357(4):370–379.
- [14] Perkins JM, Subramanian SV, Christakis Na. Social networks and health: A systematic review of sociocentric network studies in low- and middle-income countries. *Social science & medicine*. 2014 aug;125(1):60–78. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25442969>.
- [15] Coleman J, Katz E, Menzel H. The Diffusion of an Innovation among Physicians. *Sociometry*. 1957;20(4):253–270.
- [16] VanderWeele TJ. Inference for influence over multiple degrees of separation on a social network. *Statistics in medicine*. 2013 feb;32(4):591–6; discussion 597–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23341081>.
- [17] Lienert J, Marcum CS, Finney J, Reed-Tsochas F, Koehly L. Social influence on 5-year survival in a longitudinal chemotherapy ward co-presence network. *Network Science*. 2017;5(3):308–327.
- [18] Hughes JJ. Paying injection drug users to educate and recruit their peers: why participant-driven interventions are an ethical public health model. *Quality management in health care*. 1999;7(4):4–12.

- 1  
2  
3  
4  
5 [19] Ssali S, Wagner G, Tumwine C, Nannungi A, Green H. HIV Clients as Agents for Prevention: A Social  
6 Network Solution. *AIDS research and treatment*. 2012;2012:815823.  
7  
8  
9 [20] Cislighi B, Denny EK, Cisse' M, Gueye P, Shrestha B, Shrestha PN, et al. Changing Social Norms: the  
10 Importance of "Organized Diffusion" for Scaling Up Community Health Promotion and Women  
11 Empowerment Interventions. *Prevention Science*. 2019;p. 1–11.  
12  
13  
14 [21] Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden popula-  
15 tions for HIV surveillance. *Aids*. 2005;19:S67–S72.  
16  
17  
18 [22] Smith R, Rossetto K, Peterson BL. A meta-analysis of disclosure of one's HIV-positive status, stigma  
19 and social support. *AIDS care*. 2008;20(10):1266–1275.  
20  
21  
22 [23] Watters JK, Biernacki P. Targeted sampling: options for the study of hidden populations. *Social*  
23 *problems*. 1989;36(4):416–430.  
24  
25  
26 [24] Benjamin-chung J, Abedin J, Berger D, Clark A, Jimenez V, Konagaya E, et al. Spillover effects on  
27 health outcomes in low- and middle-income countries: a systematic review. *International Journal of*  
28 *Epidemiology*. 2017;(September):1–26.  
29  
30  
31 [25] Neema S, Atuyambe LM, Otolok-Tanga B, Twijukye C, Kambugu A, Thayer L, et al. Using a clinic  
32 based creativity initiative to reduce HIV related stigma at the Infectious Diseases Institute, Mulago  
33 National Referral Hospital, Uganda. *African health sciences*. 2012;12(2):231–239.  
34  
35  
36 [26] Smith Fawzi MC, Siril H, Li N, McAdam K, Ainebyona D, McAdam E, et al. Agents of Change:  
37 stepped wedge randomized controlled trial of the NAMWEZA intervention with people living with  
38 HIV and members of their social networks. *BMJ Global Health*. 2019;.  
39  
40  
41 [27] Kaaya S, Arthur G, Siril H, Lienert J, Li N, McAdam K, et al. Leadership of PLWH as Agents of  
42 Change: comparing HIV-related risk behavior of people attending ART clinics in Dar es Salaam with  
43 members of their social networks. In: 20th International AIDS Conference; 2014..  
44  
45  
46 [28] Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC medical research*  
47 *methodology*. 2006;6(1):54.  
48  
49  
50  
51  
52



- 1  
2  
3  
4  
5 [29] Zou J, Yamanaka Y, John M, Watt M, Ostermann J, Thielman N. Religion and HIV in Tanzania:  
6 influence of religious beliefs on HIV stigma, disclosure, and treatment attitudes. *BMC public health*.  
7 2009;9(1):75.  
8  
9  
10 [30] McCarthy B. Appreciative Inquiry: an alternative to behaviour management. *Dementia*.  
11 2017;16(2):249–253.  
12  
13  
14 [31] Siril HN, Kaaya SF, Fawzi MKS, Mtisi E, Somba M, Kilewo J, et al. CLINICAL outcomes and loss to  
15 follow-up among people living with HIV participating in the NAMWEZA intervention in Dar es  
16 Salaam, Tanzania: a prospective cohort study. *AIDS research and therapy*. 2017;14(1):18.  
17  
18  
19 [32] Carey MP, Schroder KE. Development and psychometric evaluation of the brief HIV Knowledge  
20 Questionnaire. *AIDS education and prevention*. 2002;14(2):172–182.  
21  
22  
23 [33] Haile BJ, Chambers JW, Garrison JL. Correlates of HIV knowledge and testing: Results of a 2003  
24 South African survey. *Journal of Black Studies*. 2007;38(2):194–208.  
25  
26  
27 [34] Aarø LE, Breivik K, Klepp KI, Kaaya S, Onya HE, Wubs A, et al. An HIV/AIDS knowledge scale for  
28 adolescents: item response theory analyses based on data from a study in South Africa and Tanzania.  
29 *Health education research*. 2011;26(2):212–224.  
30  
31  
32 [35] Peltzer K, Matseke G, Mzolo T, Majaja M. Determinants of knowledge of HIV status in South Africa:  
33 results from a population-based HIV survey. *BMC public health*. 2009;9(1):174.  
34  
35  
36 [36] Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. *Pharmaceutical Statistics: The*  
37 *Journal of Applied Statistics in the Pharmaceutical Industry*. 2009;8(1):50-61.  
38  
39  
40 [37] Easterly W. How the millennium development goals are unfair to Africa. *World development*.  
41 2009;37(1):26–35.  
42  
43  
44 [38] Barbarin OA, Khomo N. Indicators of economic status and social capital in South African townships:  
45 What do they reveal about the material and social conditions in families of poor children? *Childhood*.  
46 1997;4(2):193–222.  
47  
48  
49 [39] Smith JA, McPherson M, Smith-Lovin L. Social distance in the United States: Sex, race, religion, age,  
50  
51  
52

- 1  
2  
3 and education homophily among confidants, 1985 to 2004. *American Sociological Review*. 2014  
4 Jun;79(3):432-56.  
5  
6  
7 [40] Lesko CR, Edwards JK, Cole SR, Moore RD, Lau B. When to censor? *American journal of epidemi-*  
8 *ology*. 2017;187(3):623–632.  
9  
10  
11 [41] Hernán MA. The hazards of hazard ratios. *Epidemiology (Cambridge, Mass.)*. 2010;21(1):13.  
12  
13 [42] VanderWeele T. *Explanation in causal inference: methods for mediation and interaction*. Oxford  
14 University Press; 2015.  
15  
16 [43] Forastiere L, Mealli F, Vanderweele TJ. Identification and Estimation of Causal Mechanisms in  
17 Clustered Encouragement Designs: Disentangling Bed Nets Using Bayesian Principal Stratification  
18 Identification and Estimation of Causal Mechanisms in Clustered Encouragement Designs : Disentan-  
19 gling Be. *Journal of the American Statistical Association*. 2016;111(514):510–525. Available from:  
20 <http://dx.doi.org/10.1080/01621459.2015.1125788>.  
21  
22  
23 [44] Shalizi CR, Thomas AC. Homophily and contagion are generically confounded in observational social  
24 network studies. *SMR*. 2012;40(2):211–39.  
25  
26 [45] De La Haye K, Robins G, Mohr P, Wilson C. Homophily and contagion as explanations for weight  
27 similarities among adolescent friends. *Journal of Adolescent Health*. 2011;49(4):421–427.  
28  
29 [46] Leung MP. Treatment and spillover effects under network interference. *Review of Economics and*  
30 *Statistics*. 2016;p. 1–42.  
31  
32 [47] Vanderweele TJ, Hong G, Jones SM, Brown JL. Mediation and spillover effects in group-randomized  
33 trials: a case study of the 4Rs educational intervention. *Journal of the American Statistical Association*.  
34 2013;108(502):469–482. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23997375>.  
35  
36 [48] Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal  
37 interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological*  
38 *methods*. 2013;18(2):137.  
39  
40 [49] Latkin CA, Davey-Rothwell MA, Knowlton AR, Alexander KA, Williams CT, Boodram B. Social  
41 network approaches to recruitment, HIV prevention, medical care, and medication adherence. *Journal*  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

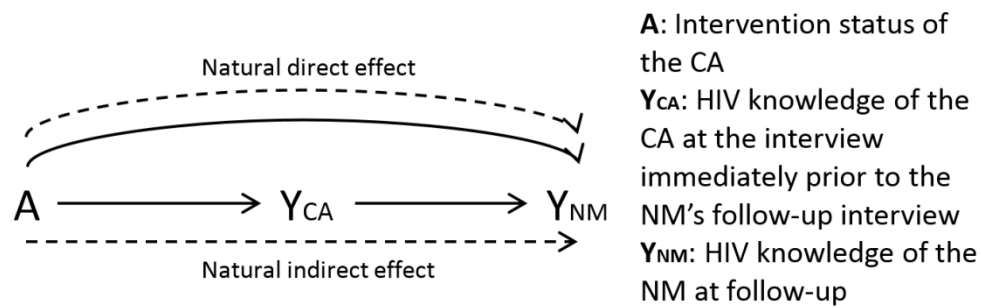
of Acquired Immune Deificiency Syndromes. 2013;63:1–9.

- [50] Galea S, Tracy M. Participation rates in epidemiologic studies. *Annals of epidemiology*. 2007;17(9):643–653.
- [51] Berkman LF. Social support, social networks, social cohesion and health. *Social work in health care*. 2000 Sep 26;31(2):3-14.
- [52] Langford CPH, Bowsher J, Maloney JP, Lillis PP. Social support: a conceptual analysis. *Journal of advanced nursing*. 1997;25(1):95–100.
- [53] Schneider JA, Zhou AN, Laumann EO. A new HIV prevention network approach: sociometric peer change agent selection. *Social science & medicine*. 2015;125:192–202.
- [54] He P, Eriksson F, Scheike TH, Zhang MJ. A proportional hazards regression model for the subdistribution with covariates-adjusted censoring weight for competing risks data. *Scandinavian Journal of Statistics*. 2016 Mar;43(1):103-22.

## 11. Figure Captions

Fig 1. Schematic of natural direct effect (NDE) and natural indirect effect (NIE). The NDE indicates the increase in NMs' HIV knowledge happens without a concomitant increase in their CA's HIV knowledge. The NIE, on the other hand, indicates that the increase in NMs' HIV knowledge is mediated by their CA's HIV knowledge increasing.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



## STROBE (Strengthening The Reporting of OBservational Studies in Epidemiology) Checklist

A checklist of items that should be included in reports of observational studies. You must report the page number in your manuscript where you consider each of the items listed in this checklist. If you have not included this information, either revise your manuscript accordingly before submitting or note N/A.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).

Section and Item	Item No.	Recommendation	Reported on Page No.
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	
<b>Introduction</b>			
Background/Rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
<b>Methods</b>			
Study Design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed  <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	

Section and Item	Item No.	Recommendation	Reported on Page No.
Data Sources/ Measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study Size	10	Explain how the study size was arrived at	
Quantitative Variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	
Statistical Methods	12	(a) Describe all statistical methods, including those used to control for confounding	
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed  <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed  <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive Data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome Data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	

Section and Item	Item No.	Recommendation	Reported on Page No.
Main Results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other Analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key Results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
<b>Other Information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Once you have completed this checklist, please save a copy and upload it as part of your submission. DO NOT include this checklist as part of the main manuscript document. It must be uploaded as a separate file.**

# BMJ Open

## Evaluating spillover of HIV knowledge from study participants to their network members in a stepped-wedge behavioral intervention in Tanzania

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-033759.R3
Article Type:	Original research
Date Submitted by the Author:	12-Jun-2020
Complete List of Authors:	Lienert, Jeffrey; University of Pennsylvania Perelman School of Medicine, Center for Health Care Innovation; Philadelphia VAMC, CHERP Fawzi, Mary C; Harvard Medical School, Global Health and Social Medicine McAdam, Keith; London School of Hygiene and Tropical Medicine, Epidemiology Kaaya, Sylvia; Muhimbili University of Health and Allied Sciences, School of Medicine Liu, Yuanyuan Todd, Jim; London School of Hygiene and Tropical Medicine, Population Health Andrew, Irene; Management and Development for Health, Onnela, Jukka Pekka; Harvard T.H. Chan School of Public Health, Biostatistics
<b>Primary Subject Heading</b>:	HIV/AIDS
Secondary Subject Heading:	Global health, Public health
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, STATISTICS & RESEARCH METHODS

SCHOLARONE™  
Manuscripts





I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3  
4  
5 Evaluating spillover of HIV knowledge from study participants to their  
6 network members in a stepped-wedge behavioral intervention in Tanzania  
7  
8

9 Jeffrey Lienert<sup>a,b,j</sup>, Mary C. Smith Fawzi<sup>c</sup>, Keith McAdam<sup>d</sup>, Sylvia Kaaya<sup>e</sup>, Yuanyuan Liu<sup>i</sup>, Jim Todd<sup>g</sup>,  
10 Irene Andrew Lema<sup>f</sup>, Jukka-Pekka Onnela<sup>h</sup>  
11

12 <sup>a</sup>Perelman School of Medicine, University of Pennsylvania <sup>b</sup>Center for Health  
13 Equity Research and Promotion, Philadelphia VA <sup>c</sup>Department of Global Health  
14 and Social Medicine, Harvard Medical School <sup>d</sup>Department of Clinical Tropical  
15 Medicine, London School of Hygiene and Tropical Medicine <sup>e</sup>Department of Psychiatry and  
16 Mental Health, Muhimbili University of Health and Allied Sciences <sup>f</sup>Management and  
17 Development for Health <sup>g</sup>Department of Population Health, London School of  
18 Hygiene and Tropical Medicine <sup>h</sup>Department of Biostatistics, Harvard T. H. Chan School of  
19 Public Health <sup>i</sup>University of Texas School of Public Health  
20 <sup>j</sup>Corresponding Author: Jeffrey Lienert, jeffrey.lienert@penncmedicine.upenn.edu  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

52 Preprint submitted to BMJ Open  
53

11 June, 2020  
54  
55  
56  
57  
58  
59  
60

## Abstract

**Objectives:** We aim to describe the social network members of participants of a behavioral intervention, and examine how the effects of the intervention may spillover among network members.

**Design:** Secondary analysis of a step-wedge randomized controlled trial.

**Setting:** Change Agents (CAs) were recruited from waiting rooms of HIV treatment facilities in Dar es Salaam, Tanzania, and their Network Members (NMs) were recruited directly by CAs.

**Participants:** We enrolled 662 CAs in an HIV behavioral intervention. They, along with 710 of their NMs, completed baseline and follow-up interviews from 2011 to 2013.

**Primary and Secondary Outcomes:** The primary outcome of this study was change in NMs' HIV knowledge, and the secondary outcome was whether the NM was lost to follow-up.

**Results:** At baseline, many characteristics were different between NMs and CAs. We found a number of NM characteristics significantly associated with follow-up of NMs, particularly female gender (OR=1.64, 95% CI: 1.02,2.63) and HIV knowledge (OR=20.0, 95% CI: 3.70,125); only one CA variable was significantly associated with NM follow-up: having a private source of water (OR=2.17, 95% CI: 1.33,3.57). The 14.2% increase in NMs' HIV knowledge was largely due to CAs feeling empowered to pass on prior knowledge, rather than transmitting new knowledge to their NMs.

**Conclusions:** Characteristics of social network members of persons living with HIV PLH may play a role in study retention. Additionally, the HIV knowledge of these NMs increased largely as a function of CA participation in the intervention, suggesting that intervening among highly-connected individuals may maximize benefits to the potential population for whom spillover can occur.

*Keywords:*

mediation, natural direct effect, natural indirect effect, follow-up, retention, hidden population

## 1. Strengths and Limitations of This Study

- The recruitment method, focused on social network members of persons living with HIV, allowed us to access an at-risk population typically not easily-accessible via other means, such as direct approach by the study team rather than someone known to the participant.
- The study design of the trial in which this study is nested was an ideal situation in which to employ a mediation analysis that informs our understanding of how this and similar intervention effects may spillover in a population.
- The greater-than-ideal dropout rate of NMs was both a strength in that it allowed us to examine factors associated with dropout, but also a limitation, in that the potential of differential dropout by unmeasured factors may have biased some of our results.

## 2. Introduction

Despite recent decreases in mortality, HIV/AIDS is currently the ninth leading cause of years of life lost (YLL) globally [1]. Research into eliminating HIV has focused on two fronts: biological and behavioral, which combined have decreased the years of life lost due to HIV by an estimated 51% [1]. Behavioral interventions seek to curb transmission of HIV through reduction of risk behavior including safer sexual practices and reducing injection drug use [2, 3]. Behavioral change also impacts the effectiveness of the biological methods, given that lower adherence is known to often compromise their effectiveness [4]. However, neither biological nor behavioral methods are fully effective without the knowledge of these treatments and preventative behaviors. It is therefore important to increase HIV-related knowledge as a necessary, but not sufficient, step towards achieving the 90-90-90 treatment goals (90% diagnosis, 90% antiretroviral therapy, and 90% viral suppression among the treated) and ultimately eliminate HIV [5].

The importance of adequate HIV knowledge has been recognized for a period of time, as it is often a prerequisite for behavioral change [6]. It should be noted that increased knowledge does not necessarily translate into behavior change, as other factors such as motivation and skills also play a role [6]. Despite this, researchers conducted trials among Persons Living with HIV (PLH) which led to increased HIV knowledge [7, 8]. These interventions were group discussions and teaching led by nurses, and motivational interviewing, respectively. Researchers also found that the intervention reduced HIV-risk behaviors

1  
2  
3 concomitantly with an increase in HIV knowledge [8].  
4

5 Knowledge gained by participants in these trials can also be freely shared with members of their social  
6 network, in what is known as a spillover effect [9]. Specifically, a spillover effect (also known as an in-  
7 direct or disseminated effect) is one person's exposure affecting another's outcome [10]. In the context of a  
8 social network spillover intervention, it means a change to someone's behavior who did not receive the  
9 intervention because they were socially connected to someone who did receive the intervention. This is dis-  
10 tinct from what is sometimes called behavioral spillover, where change's in a person's behavior affects other  
11 behaviors of that same person [11]. For injection drug users, HIV prevention educational interventions were  
12 demonstrated to have spillover effects of HIV prevention education, and subsequent reduced rates of risky  
13 behaviors [12]. What remains unknown, however, is whether or not a spillover effect exists for HIV  
14 knowledge during and after an intervention in other populations, particularly sub-Saharan Africa. In other  
15 words, we aim to determine that if an intervention increases someone's HIV knowledge, how members of  
16 their social networks also increase their HIV knowledge.  
17  
18  
19  
20  
21  
22  
23  
24

25 New knowledge can come from a variety of sources, one of the most important of which is a person's  
26 social network [13, 14]. Social networks are of particular import because new knowledge can lead to  
27 cascades of behavior change, where people subsequently educate those in their social network, in what is  
28 known as social influence [15, 16, 17]. This has been examined in participant-driven interventions, where  
29 initially-recruited participants educate members of their social network one-on-one [18]. Characteristics  
30 such as knowing the HIV status of network members has been shown to be the most important predictor of  
31 engaging in prevention advocacy [19]. Work on diffusion through social networks, how a belief or behavior  
32 can be "contagious" within a network, has shown that spreading intervention effects beyond the initial study  
33 population can improve the cost-effectiveness of these interventions [20]. These findings imply that certain  
34 aspects of knowledge or behavior may spread more or less efficiently through networks comprising  
35 individuals with specific characteristics, which may need to be accounted for in network interventions. For  
36 instance, networks comprising many at-risk individuals who are HIV-negative may not be as receptive to  
37 change in behaviors as networks comprising a mix of those with and without HIV. Additionally, the  
38 networks of PLH are often difficult networks to ascertain [21], due to the continued stigma of HIV and AIDS  
39 in many settings [22]. Because of this, if a PLH or person at-risk of acquiring HIV does not want to  
40 participate in an investigator-initiated intervention, there is little recourse other than information transmitted  
41 via social networks, or targeted sampling techniques which are not always effective (e.g. Respondent Driven  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53

1  
2  
3 Sampling) [23]. This is particularly important in low- and middle-income countries, as a recent systematic  
4 review found only 54 studies researching spillover effects in these settings (out of approximately 750) [24].  
5  
6 Therefore, understanding exactly how information spreads from participants in an intervention to members  
7 of their social network, who may be largely inaccessible via other means, is important for reaching the  
8 greatest number of people about HIV prevention. Understanding what makes these persons different from  
9 those who receive the intervention themselves is important, as it may point to ways in which to increase  
10 enrollment of these populations.  
11  
12

13  
14  
15 Based on the above gaps in the literature, we conducted a study on network members of PLH enrolled  
16 in a behavior change intervention [25]. The trial recruited PLH to serve as Change Agents (CAs) and to  
17 reach out to their social network members (NMs) about knowledge of HIV and safer sexual practices [26].  
18 Our goals in this study were threefold: 1) to understand how the NMs differed from their CAs, 2) understand  
19 correlates of dropout for the NMs, and 3) to understand by what method HIV knowledge transferred to the  
20 NMs from the CAs. Understanding how the information and behaviors are shared within social networks  
21 will allow HIV researchers and others to take advantage of this knowledge and improve upon prevention  
22 interventions in the future.  
23  
24  
25  
26  
27  
28

### 29 **3. Methods**

#### 30 *3.1. Study Population*

31  
32 We analyzed social network data from the Agents of Change trial [27], which was a stepped-wedge  
33 randomized controlled trial [28] that enrolled PLH to become Change Agents (CAs) by informing members  
34 of their social network (NMs) about knowledge of HIV and safer sexual practices. Here, we define CAs  
35 based on the *potential* for PLH to become so - by self-selecting into the study, PLH identify themselves as  
36 potential CAs. The NAMWEZA intervention is then designed to foster a CA's ability to truly act as a Change  
37 Agent, rather than in name only.  
38  
39  
40  
41  
42  
43

44 CAs were recruited from the waiting rooms of HIV care and treatment centers in Dar es Salaam, Tanza-  
45 nia, and we received written consent from each CA. Participants were given a baseline questionnaire and  
46 were randomized to one of three waves in which to receive the intervention. At baseline, participants were  
47 also asked to recruit up to three members of their social networks who they felt were at particularly high risk  
48 of contracting or spreading HIV. We obtained written consent from these nominated network members (NMs).  
49 NMs could be either HIV positive or negative, and they were given a baseline survey. The NM was only  
50  
51  
52  
53

1  
2  
3 aware the CA was a participant in the intervention if the CA shared this information with them, which  
4 many did not due to HIV-related stigma [29]. Each CA therefore formed a CA-NM dyad with each NM they  
5 recruited, and if they recruited more than one NM, formed a set of CA-NM dyads with a common CA.  
6  
7

8 As fits a stepped-wedge RCT, all CAs eventually received the intervention, but were randomized to  
9 *when* they received it. These waves each lasted 12 weeks, at which point the next wave began and another  
10 group of CAs received the intervention. Within each wave, the intervention comprised 10 weekly structured  
11 sessions aimed at empowering PLH to become HIV prevention change agents in their communities. The  
12 sessions aimed to increase psychosocial and communication skills through an Appreciative Inquiry approach  
13 [30]. Within one month of each wave of the intervention, CAs were given follow-up surveys. Across all  
14 waves, the interventions lasted from November 2010 to January 2014, and the final interviews were  
15 conducted in March 2014. For more information on the study design, we direct interested readers to Smith-  
16 Fawzi et al., 2019 [26].  
17  
18  
19  
20  
21  
22

23 The NMs did not receive any intervention at any point during the study. Rather, their intervention status  
24 flowed from the intervention status of their CAs. Due to operational difficulties, each NM was surveyed two  
25 times during the study: baseline and after the first wave, rather than baseline and one after each wave. In  
26 this way, all demographic and contextual variables were measured at baseline. We use this interview of NMs  
27 as the division between “exposed” and “unexposed”. At the time of an NM’s follow-up interview, their  
28 respective CA may or may not have undergone the intervention. Therefore, the NMs were divided into  
29 “exposed” (N=381) and “unexposed” (N=329) groups based on whether their respective CA had completed  
30 their intervention at the time of the NM’s follow-up interview.  
31  
32  
33  
34  
35  
36

37 Inclusion criteria for the follow-up analyses were for CAs who completed at least one follow-up inter-  
38 view and had at least one NM who completed at least one follow-up questionnaire. In this way, outcomes  
39 could be computed for the CA and NM of each dyad. NMs approached by a CA, but who did not participate  
40 in the study were not recorded since it was not feasible to obtain this information from study participants  
41 themselves. As we lost some CAs and NMs to follow up, we completed our analyses without their data,  
42 assuming it to be Missing Completely at Random (MCAR). During this study, there was little loss-to-follow-  
43 up among the CAs (< 10%), but much higher loss among the NMs (36.8%) [31]. Given an NM or CA was  
44 not lost to follow-up, complete information was available on all additional variables, including exposure,  
45 outcome, and covariates. In sum, our sample comprises 662 CAs and 710 NMs, meaning each CA recruited  
46 1.07 NMs on average out of a possible 3, and 44 CAs nominated at least two NMs.  
47  
48  
49  
50  
51  
52

### 3.2. *HIV knowledge*

To assess HIV knowledge of CAs and NMs, we used the brief HIV knowledge questionnaire [32]. This scale comprises 18 questions, which focus on an individual's knowledge of how HIV can spread and other characteristics of the virus and AIDS. This knowledge is crucial for an individual's subsequent safe-sex practices to reduce the risk of transmission. The original population comprised three different groups: two groups of low-income women, and one of women and men receiving psychiatric treatment. In these populations, the measure had a Chronbach's alpha of 0.78. This instrument has been used previously in sub-Saharan Africa and has demonstrated reasonable reliability of 0.75 Cronbach's alpha in South Africa among a convenience sample of 429 members of the African Methodist Episcopal church [33]. It has also been translated to Swahili, with only minor differential item functioning [34]. This indicates that the measure performs adequately in other, similar populations, relevant to this work. Here, the Cronbach's alpha was 0.71. Because the average baseline score of HIV knowledge was 0.80, with a preponderance of scores of 1.00, the main indicator of knowledge was not normally distributed, and therefore a continuous predictor was not ideal [35,36]. We therefore summarize this measure as "Complete HIV Knowledge", a dichotomous variable for whether the participant correctly answered all questions.

### 3.3. *Demographic and contextual variables*

In addition to HIV knowledge, we are interested in a set of demographic and contextual variables that may help explain some of the trends observed. In terms of demographic variables, we include age, sex, employment status, marital status, and self-identified HIV status. We also include education, which we dichotomize at 7 years or more - a level commensurate with elementary education. We selected this cutoff because it coincides with the millennium development goal (MDG) of increasing primary education completion [37].

Contextually, we included two additional variables: having a private source of water, and the number of persons sleeping in the participant's home. The first was also based on a MDG, and indicates participants with access to safe drinking water [37]. This is a proxy for the economic security of the participant. The number of persons sleeping in the participant's home is also a non-monetary indicator of their material and social resources [38]. These variables combined give a more thorough picture of the participant's economic status than employment alone.



### 3.4. Statistical Analysis

The first analysis was to assess homophily, defined as the extent to which CAs and NMs were similar to one another in terms of a number of sociodemographic characteristics and HIV-related risk factors. 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. Because CAs and NMs self-selected into the study and were not randomized to CA/NM status, we do not *a priori* expect them to be completely similar. In addition, multiple NMs could share a CA and would therefore not be independent due to the shared variation and latent characteristics of having the same CA. 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. For continuous variables, the difference between the CA and NM was calculated, and for categorical variables, whether the CA and NM were concordant or discordant was recorded. 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. We then examined the percentile of the observed difference relative to the permuted differences [39]. Analyses were run using R v3.1.1.

To accomplish our second aim of understanding what was associated with NMs completing their follow-up interview, we fit a log-binomial regression to determine predictors of follow-up. In this regression, we use all the variables listed above, as well as whether the CA remained in the study for its full duration. In order to examine the association between these same variables and the time of follow-up, we also employed a Cox regression. In the Cox regression, start of follow-up was defined as the time at which the NM completed their baseline interview. The outcome here was whether the NM completed a follow-up interview. NMs who were lost to follow-up were censored at the time of their latest interview [40,41].

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 passing it on, or the CAs being empowered by the intervention to pass on existing knowledge. As the wedge in which the CA received the *NAMWEZA* intervention was randomized, we treat the time at which an NM was potentially exposed to *NAMWEZA* through their CA as similarly randomized. This randomization scheme allowed us to explore the spillover effect of CAs' HIV knowledge onto their respective NMs via a mediation analysis. These pathways represent

1  
2  
3 different types of spillover effects: the exposure or outcome of one person affecting the outcome of another  
4 person.  
5

6 As shown by VanderWeele et al (2015), social network spillover effects in the case of dyadic  
7 relationships can be broken down into concepts from mediation analysis: direct and indirect effects (Figure  
8 1) [42]. This method has since been used for novel evaluations of spillover effects [43,44]. Although  
9 previous studies showed that this same type of analysis cannot be done on full network data, the data in this  
10 study consisted of only dyads, the CAs and their NMs, so in this case the method is appropriate [45, 46].  
11  
12  
13  
14

15 The method parses social influence into direct and indirect effects. The natural indirect effect (NIE) is  
16 the effect by which the intervention changes the CA's HIV knowledge, which then changes the NM's HIV  
17 knowledge. The NIE is similar to the effect observed in many participant-driven interventions: an initial  
18 participant receives the intervention, increasing their knowledge, and they subsequently pass their increased  
19 knowledge to members of their social network [18]. The natural direct effect (NDE) is the effect of receiving  
20 the intervention has on an NMs outcome, irrespective of the CA's outcome (in this case HIV knowledge).  
21 For instance, this could occur if CAs begin with good knowledge of HIV, and the intervention empowers  
22 them to convey knowledge they already had to their NMs. Although the intervention does not increase their  
23 HIV knowledge, it is still useful to the CAs, as it empowers them to become CAs in the first place. In order  
24 to estimate these effects, the published SAS macro developed by VanderWeele and colleagues was used,  
25 and analyses were run using SAS v9.2 (SAS Institute, Cary, NC) [47,48].  
26  
27  
28  
29  
30  
31  
32  
33

34 One assumption of this analysis is that the dyads are independent, which is violated here; if a CA  
35 recruited more than one NM, the multiple CA-NM dyads involving the same CA would not be independent.  
36 To address this, we performed the analysis after randomly removing NMs until each CA had only a single  
37 NM. This resulted in removing 48 NMs, just 6.7% of the population. We found that the point estimates were  
38 nearly identical, but that the confidence intervals were slightly larger due to the reduced sample size. No  
39 coefficients changed from significant to non-significant in this analysis (data not shown).  
40  
41  
42  
43

44 A second, related assumption is that of *partial interference*, that the effects in one cluster does not affect  
45 another cluster - here, one CA-NM dyad affecting another [49]. This could occur if two NMs of different  
46 CAs happen to know one-another outside of the study, one has a CA who was randomized to an earlier  
47 wedge, and shares what they know of it with the otherwise-unexposed NM. However, due to the large size  
48 of Dar es Salaam, and the number of HIV treatment clinics in which recruitment occurred, we expect few  
49 CAs or NMs to know one another outside of the study (other than NMs knowing the CA who recruited  
50  
51  
52  
53

them), limiting the potential for partial interference.

The data are not publicly available due to the sensitive nature of HIV infection status.

### 3.5. Patient and Public Involvement

Patients were not involved in the design of this study. As part of the recruitment process, patients were instructed to recruit members of their social networks who they felt were at high risk of HIV infection - in this way participants could guide the recruitment of NMs to those most at-risk. We currently have no plans to disseminate the results of this study to participants.

## 4. Results

The *NAMWEZA* study recruited 662 Change agents (CA), of whom 453 (68.4%) completed at least one follow-up questionnaire. These 662 CAs in turn recruited 710 network members (NM) who took the baseline questionnaire. Of the 710 NMs, 449 (63.2%) also completed a follow-up questionnaire (Table 1). At baseline, CAs were on average older than their NMs. More of the NMs were employed than their CAs (69.3% vs. 55.0%,  $p < 0.001$ ), but were less likely to have at least 7 years of education (52.0% vs. 52.3%,  $p < 0.001$ ). Only 12.3% of NMs were HIV-positive, compared to *all* CAs ( $p < 0.001$ ). Complete data was obtained at baseline for all CAs and NMs.

Characteristic	Number of NMs (%) or Mean (SD) (N=710)	Number of CAs (%) or Mean (SD) (N=662)	P-value
Age	33.0 (11.1)	38.9 (9.7)	<0.001
Female	380 (53.7%)	349 (53.9%)	0.89
Employed	490 (69.3%)	356 (55.0%)	<0.001
At least 7 years education	369 (52.0%)	584 (82.3%)	<0.001
Complete HIV knowledge	638 (89.9%)	598 (90.4%)	0.65
Persons sleeping in home	5.00 (5.77)	3.87 (3.72)	<0.001
Married	373 (52.7%)	338 (51.1%)	0.56
HIV Positive	87 (12.3%)	662 (100%)	N/A
Private source of water	309 (43.7%)	263 (39.7%)	0.19

Table 1: Demographic characteristics at baseline, with the results of a permutation test for homophily. Specifically, for each CA- NM dyad, either a difference (for continuous variables) or concordance (for dichotomous variables) is calculated. For example, if a CA was 39 years old, and their NM was 25 years old, the difference would be 14 years old. If a CA was male and their NM was Female, the pair would be discordant for sex. CA-NM pairs were then randomly reshuffled, the edge-wise characteristics recalculated, and the observed difference compared to the distribution of randomized differences.

1  
2  
3  
4 Logistic regression showed that characteristics of both the CAs and the NMs significantly predicted loss  
5 to follow-up (Table 2). The NM being female (RR=1.44, 95% CI: 1.05,1.97), having complete HIV  
6 knowledge (RR=10, 95% CI: 2.33,42), being employed (RR=1.43, 95% CI: 1.08,1.89), and being married  
7 (RR=1.55, 95% CI: 1.03,2.33) were all significant predictors of increased odds of completing a follow-up  
8 interview. Each additional person sleeping in the home of the NM per room used for sleeping reduced the  
9 odds of follow-up (RR=0.85, 95% CI: 0.74,0.98) as did the NM living with HIV (RR=0.40, 95% CI:  
10 0.17,0.96). CA having a private water source was significantly associated with increased odds of the NM  
11 being followed-up (RR=2.07, 95% CI: 1.25,3.42), even after controlling for the NM having a private source  
12 of water. This was the only CA-specific variable that significantly predicted an NM's follow-up.  
13  
14  
15  
16  
17  
18

19 The Cox proportional hazard model showed similar results to the logistic regression, but with fewer  
20 significant results. Only the NM being married (HR=1.28, 95% CI: 1.01,1.64) and the NM living with HIV  
21 (HR=0.71, 95% CI: 0.51,0.99) significantly predicted the time-to-follow-up of the NMs. Additionally, all  
22 of the hazard ratios were closer to the null relative to the corresponding odds ratios.  
23  
24  
25

26 We assess the mechanism of spillover effects on HIV knowledge, the main effect of the intervention on  
27 the NMs. The analysis showed that NMs who were exposed to intervention via their CA had an increase in  
28 HIV knowledge of 12.9% on average (95% CI: 0.06, 0.20) from a baseline percentage of 78% [26]. This  
29 was broken down into a natural indirect effect (NIE) of 0.6% (95% CI: -0.06%,2.0%), which is the effect  
30 the intervention had via the CA's HIV knowledge, and a natural direct effect (NDE) of 12.3% (95% CI:  
31 6.1%,19.3%), which is the effect the CA's participation had on the NM's HIV knowledge, irrespective of  
32 the CA's HIV knowledge. In other words, of the 12.9% increase in NM's having complete HIV knowledge,  
33 12.3% occurred without a concomitant increase in their CA's HIV knowledge. In other words, their HIV  
34 knowledge increase was not mediated by the increase in HIV knowledge of their CA. This did not change  
35 when we used only one NM per CA.  
36  
37  
38  
39  
40  
41

## 42 **5. Discussion**

43  
44 Based on results from the NAMWEZA trial in Tanzania, we have shown a number of novel findings  
45 regarding the network members of HIV intervention participants, correlates of their retention in the study,  
46 and the evidence for spillover of HIV knowledge from CAs to NMs. These findings can inform the design  
47 of interventions in the future to maximally enroll and retain participants and ensure that intervention  
48 information is transmitted from the study participants to members of their social networks.  
49  
50  
51

Characteristic (N=459)	Adjusted RR (95% CI)	Adjusted HR (95% CI)
Characteristics of NMs		
Gender (Female)	1.44* (1.05,1.97)	1.18 (0.94,1.50)
Difference in age of NM and CA (per year)	1.01 (0.99,1.03)	1.01 (0.995,1.02)
Age of NM	0.98 (0.95,1.02)	0.99 (0.97,1.04)
Complete HIV knowledge (vs. none)	10* (2.33,42)	2.20 (0.97,5.01)
Employed	1.43* (1.08,1.89)	1.15 (0.89,1.50)
Each additional person sleeping in home per room used for sleeping	0.85* (0.74,0.98)	0.92 (0.83,1.01)
Married	1.55* (1.03,2.33)	1.28* (1.01,1.64)
Living with HIV	0.40* (0.17,0.96)	0.71* (0.51,0.99)
Having a private source of water	0.89 (0.59,1.34)	0.97 (0.77,1.22)
Characteristics of CAs		
Gender (Female)	1.27 (0.76,2.08)	1.07 (0.84,1.37)
Complete HIV knowledge (vs. none)	0.36 (0.07,2.04)	0.64 (0.29,1.43)
Having a private source of water	2.07* (1.25,3.42)	0.97 (0.77,1.22)
Being employed	1.54 (0.95,2.50)	1.14 (0.90,1.43)
Being married	1.25 (0.77,2.04)	1.11 (0.88,1.43)
CA Lost to follow-up	1.06 (0.84,1.33)	1.02 (0.72,1.43)

Table 2: Results of multivariate log-binomial regression and Cox-proportional hazard models on the dichotomous outcomes of whether the NM completed a follow-up questionnaire, and the continuous outcome of time-to-completion of follow-up questionnaire, respectively. “\*” indicates significant at the  $p < 0.05$  level.

We found several significant differences between CAs and NMs they recruited. On average, CAs were older, less likely to be employed, and more educated. All of this suggests that by having CAs recruit from their social network, we were able to recruit a set of social network members different in multiple ways from those found in the waiting rooms in HIV treatment clinics (unsurprising, given that many NMs were not HIV-positive) [9]. This may in part be due to the fact that when recruiting NMs, CAs were instructed to recruit those who may be at particular risk of contracting or spreading HIV. Therefore, CAs likely did not pick random members of their social network, but those who were at high risk. This finding corroborates the claim by Latkin et al. (2013) that sampling via enumeration of network members by initial participants recruits a more diverse sample, as our sample of NMs was not composed of only those who were HIV-positive, but also many who were HIV-negative [50]. Additionally, the lower likelihood of employment of the CAs indicates they may have had more time for participating in the intervention [51]. This means that future studies may need to tailor their interventions to work with the schedules of employed persons to increase participation.

Although the study design potentially accessed a separate slice of the population than other methods vis-a-vis the population of those at risk for contracting HIV, it may have been at least partially responsible for the large loss of follow-up of the NMs. It is important to note that more NM characteristics than CA

1  
2  
3 characteristics were significantly predictive of follow-up in both the logistic and Cox proportional hazards  
4 models. The one CA characteristic which did predict NMs completing a follow-up interview, having a  
5 private source of water, was a proxy for CA's socioeconomic status. This may have been because CAs with  
6 greater resources may have had more time available to pass information to their NMs, retaining the NM's  
7 interest longer [52]. We also note that less HIV knowledge and living with HIV predicted reduced likelihood  
8 of completing a follow-up questionnaire among NMs, which means that those who might have benefited  
9 most from spillover of the intervention were more likely to discontinue their involvement. This does not  
10 mean that they did not receive any spillover, only that it was not recorded. Our estimate of the magnitude of  
11 the spillover may therefore be biased towards the null. This is also problematic more generally for  
12 interventions of this nature as the very people the intervention aims to benefit may not stay with the program.  
13  
14

15  
16 The HIV knowledge gain experienced by the NMs was largely due to the NDE; i.e. knowledge spilled-  
17 over as a result of the CAs participating in NAMWEZA, not because the CAs' HIV knowledge increased,  
18 and they passed this new knowledge to their NM(s). This is most likely because CAs had a high average  
19 HIV knowledge score at baseline (80%), so there was less room for improvements in knowledge. Following  
20 the intervention, CAs' likelihood of complete HIV knowledge did not significantly increase. Therefore, what  
21 prompted the increased knowledge of the NMs was the CA becoming empowered through the intervention  
22 to pass on their existing knowledge to their NM [53].  
23  
24

25  
26 This finding is important for future interventions: spillover effects of this intervention will likely carry  
27 over only to those directly-connected to the CAs (opposed to spreading indefinitely in a snowball effect), as  
28 the most important thing in transmitting HIV knowledge is receiving the intervention, which only CAs  
29 receive. Because NMs do not actually receive the intervention, it is unlikely that they would feel empowered  
30 to become CAs themselves (particularly because they will not receive the *NAMWEZA* sessions), thus limiting  
31 the spread of the intervention. For CAs to increase the HIV knowledge of their NMs, they would only need  
32 to become empowered to share their information. However, for the NMs to increase the HIV knowledge of  
33 their own NMs, they would need to both increase their HIV knowledge, and become empowered to share it.  
34 This is less likely to happen than just becoming empowered, and so it is unlikely, though not impossible, that  
35 this effect would continue to spread in the population. This may give insight into how to design interventions  
36 in the future; if one wants to maximize the number of people who benefit, choosing CAs who form many  
37 bridging ties in the community would maximize the potential number of links by which spillover can occur  
38 [54]. Alternatively, interventions can be designed to be self-propagating; if CAs are empowered to deliver  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

1  
2  
3 the intervention to others, changing their own NMs into future CAs, those new NMs-turned-CAs could then  
4 deliver the intervention to a second set of NMs, again empowering them to become CAs.  
5

6  
7 Our study is not without limitations. First, there was considerable loss to follow-up of the NMs. Al-  
8 though this actually informed our analysis of the correlates of loss to follow-up, it meant that our analysis  
9 of spillover effects may be biased. Even though the exposure was randomized, the loss to follow-up can  
10 result in selection bias if the NMs who left the study were systematically different from those who remained.  
11 As we show, the NMs who dropped out were those who would have benefited the most from the intervention  
12 because they were more vulnerable and at higher risk than those who did not drop out. Because they are  
13 likely to have greater benefits from any spillover effects, we would expect this non-random loss to follow-up  
14 to result in an underestimate of the impact of the intervention in the network members. Adjusting for  
15 censoring weights may ameliorate this issue [55]. Second, our use of Hazard Ratios (HRs) has important  
16 limitations: they are subject to selection bias, are sensitive to study period, and only provide one estimate  
17 during the study [41]. Any of these limitations could affect this analysis, hence our use of logistic regression  
18 as a primary analysis. However, they remain useful as a sensitivity analysis. Third, our data did not perfectly  
19 fit the requirements of the causal mediation analysis: they were not entirely independent, since multiple NM-  
20 CA dyads shared a CA. However, when we randomly removed dyads until there were no repeated CAs, the  
21 results were qualitatively very similar, indicating that lack of independence did not unduly affect our results.  
22 Fourth, although we were able to tease apart the direct and indirect effects, we are unable to determine the  
23 mechanism of the natural direct effect; the data do not allow us to specify whether NMs increased knowledge  
24 through speaking to their knowledgeable CA, through researching HIV on their own, or some other  
25 mechanism. Future work will have to be done to examine these different pathways.  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39

## 40 **6. Conclusions**

41  
42 These results have implications for the potential scale-up of the *NAMWEZA* intervention, as well as  
43 future studies and interventions that focus on behavioral interventions in social networks. First, our findings  
44 of minimal similarity between CAs and their NMs indicate that this recruitment method allows us to enroll  
45 participants from portions of the population that are not represented by the CAs alone. Coupling this with a  
46 deeper understanding of the mental heuristics CAs used to select NMs (e.g. did CAs mentally search their  
47 close or peripheral network for those at-risk of HIV), may lead to different strategies for recruitment and  
48 retention, leading to stronger effects of behavioral interventions. The mediation analysis presents a  
49  
50  
51  
52  
53

1  
2  
3 compelling picture of how best to ensure the benefits of interventions reach as many people beyond the  
4 study participants as possible. Participation of CAs in the intervention resulted in positive effects on their  
5 immediate network members' HIV knowledge regardless of how the CAs responded to the intervention.  
6 While improvement in HIV knowledge may not necessarily translate to increased safe sex behavior, it can  
7 be seen as a rate-limiting step towards reducing HIV transmission since adequate knowledge is usually  
8 necessary for reducing risk behavior. Future work should examine the exact mechanisms of spillover  
9 discussed here, as that is an important clarification that could benefit future studies. Specifically, a similar  
10 setup to the study here combined with semi-structured interviews with CAs and NMs about their interactions  
11 with one another would help elucidate exactly how NMs increased their HIV knowledge (this population  
12 would likely need to comprise CAs who have disclosed their HIV status in order to prevent accidental  
13 disclosure). Future work should also examine whether increases in HIV knowledge translate to changes in  
14 behaviors which may increase one's risk of contracting HIV, particularly in Sub-Saharan Africa. The results  
15 presented herein may inform approaches for increasing participation and potentially conferring greater  
16 benefits related to spillover effects in future HIV behavioral interventions.  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26

## 27 **7. Acknowledgements**

28  
29 The authors would like to thank the study participants and staff for their time, hard work, and dedication.  
30 We would also like to thank Guy Harling for their comments on the manuscript.  
31  
32  
33

## 34 **8. Contributorship**

35  
36 JL, MCSF, and JPO designed the analysis plan and drafted the initial paper. JPL, MCSF, KM, SK, IAL and JT  
37 conducted the underlying randomized trial, and designed it to incorporate social networks data and spillover information;  
38 these same authors provided feedback on the proposed analysis plan. JL and YL conducted analyses for the paper. All  
39 authors helped revise the drafts of the paper and approved the final manuscript.  
40  
41  
42  
43  
44  
45

## 46 **9. Compliance with ethical standards**

47  
48 **Funding:** The authors received funding from the CDC via grant TZ/UG.08.0147, the President's Emer-  
49 gency Plan for AIDS Relief - Tanzania, and the UTHealth Innovation for Cancer Prevention Research Train-  
50 ing Program Pre-doctoral Fellowship (Cancer Prevention and Research Institute of Texas grant RP160015).  
51  
52



1  
2  
3 **Conflict of interest:** The authors declare that they have no conflict of interest.  
4

5 **Ethical approval:** All procedures performed in studies involving human participants were in accordance  
6 with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki  
7 declaration and its later amendments or comparable ethical standards. IRB approval was obtained from  
8 Harvard Medical School, the Muhimbili University of Health and Allied Sciences, and the National Institute  
9 for Medical Research (NIMR) in Tanzania. Particular attention was paid to eliminate the risk of inadvertent  
10 disclosure of CA HIV status to NMs.  
11

12 **Informed consent:** Informed consent was obtained from all individual participants included in the study.  
13

14 **Disclaimer:** The content is solely the responsibility of the authors and does not necessarily represent the  
15 official views of the Cancer Prevention and Research Institute of Texas.  
16

17 **Data sharing statement:** There are no additional data available. As the data contain enough information  
18 to potentially-uniquely identify specific participants, in the sensitive context of HIV, we have chosen not to  
19 make the data available.  
20  
21

## 22 **10. References**

- 23  
24  
25  
26  
27  
28 [1] Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and  
29 national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the  
30 Global Burden of Disease Study 2016. *The Lancet*. 2017;390(10100):1151–1210.  
31  
32  
33 [2] Crepaz N, Tungol-Ashmon MV, Higa DH, Vosburgh W, Mullins MM, Barham T, et al. A systematic  
34 review of interventions for reducing HIV risk behaviors among people living with HIV in the  
35 United States, 1988-2012. *AIDS*. 2014 mar;28(5):633–56. Available from:  
36 <http://www.ncbi.nlm.nih.gov/pubmed/24983541>.  
37  
38  
39 [3] Arnold EA, Kegeles SM, Pollack LM, Neilands TB, Cornwell SM, Stewart WR, et al. A Randomized  
40 Controlled Trial to Reduce HIV-Related Risk in African American Men Who Have Sex with Men and  
41 Women: the Bruthas Project. *Prevention Science*. 2018;p. 1–11.  
42  
43  
44 [4] Garcia de Olalla P, Knobel H, Carmona A, Guelar A, Lopez-Colomes J, Cayla J. Impact of adherence  
45 and highly active retroviral therapy on survival in HIV-infected patients. *Journal of Acquired Immune  
46 Deiciency Syndromes*. 2002;30(1):105–10.  
47  
48  
49  
50  
51

- 1  
2  
3 [5] Joint United Nations Programme on HIV/AIDS and Joint United Nations Programme on HIV/Aids and  
4 others. 90-90-90: an ambitious treatment target to help end the AIDS epidemic. Geneva: UNAIDS.  
5 2014;.  
6  
7  
8  
9 [6] Fisher JD, Fisher WA. Changing AIDS-risk behavior. *Psychological bulletin*. 1992;111(3):455.  
10  
11 [7] McNeill C, George N, Glover R. An Evaluation of Sisters Informing Healing Living Empowering:  
12 Increasing HIV Knowledge Among African American Adolescent Females Using an Evidence-Based  
13 HIV Prevention Intervention. *Journal of Doctoral Nursing Practice*. 2017;10(1):4–10.  
14  
15 [8] Carey MP, Braaten LS, Maisto SA, Gleason JR, Forsyth AD, Durant LE, et al. Using information,  
16 motivational enhancement, and skills training to reduce the risk of HIV infection for low-income urban  
17 women: A second randomized clinical trial. *Health Psychology*. 2000;19(1):3.  
18  
19 [9] Sinclair B, McConnell M, Green DP. Detecting spillover effects: Design and analysis of multilevel  
20 experiments. *American Journal of Political Science*. 2012;56(4):1055–1069.  
21  
22 [10] Benjamin-Chung J, Arnold BF, Berger D, Luby SP, Miguel E, Colford Jr JM, et al. Spillover effects in  
23 epidemiology: parameters, study designs and methodological considerations. *International journal of  
24 epidemiology*. 2017;47(1):332–347.  
25  
26 [11] Truelove HB, Carrico AR, Weber EU, Raimi KT, Vandenberg MP. Positive and negative spillover of  
27 pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental  
28 Change*. 2014;29:127–138.  
29  
30 [12] Ghosh D, Krishnan A, Gibson B, Brown SE, Latkin CA, Altice FL. Social network strategies to address  
31 HIV prevention and treatment continuum of care among at-risk and HIV-infected substance users: a  
32 systematic scoping review. *AIDS and Behavior*. 2017;21(4):1183–1207.  
33  
34 [13] Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *The New  
35 England journal of medicine*. 2007;357(4):370–379.  
36  
37 [14] Perkins JM, Subramanian SV, Christakis Na. Social networks and health: A systematic review of  
38 sociocentric network studies in low- and middle-income countries. *Social science & medicine*. 2014  
39 aug;125(1):60–78. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25442969>.  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

- 1  
2  
3 [15] Coleman J, Katz E, Menzel H. The Diffusion of an Innovation among Physicians. *Sociometry*.  
4 1957;20(4):253–270.  
5  
6  
7 [16] VanderWeele TJ. Inference for influence over multiple degrees of separation on a social network.  
8 *Statistics in medicine*. 2013 feb;32(4):591–6; discussion 597–9. Available from:  
9 <http://www.ncbi.nlm.nih.gov/pubmed/23341081>.  
10  
11  
12 [17] Lienert J, Marcum CS, Finney J, Reed-Tsochas F, Koehly L. Social influence on 5-year survival in a  
13 longitudinal chemotherapy ward co-presence network. *Network Science*. 2017;5(3):308–327.  
14  
15 [18] Hughes JJ. Paying injection drug users to educate and recruit their peers: why participant-driven  
16 interventions are an ethical public health model. *Quality management in health care*. 1999;7(4):4–12.  
17  
18 [19] Ssali S, Wagner G, Tumwine C, Nannungi A, Green H. HIV Clients as Agents for Prevention: A Social  
19 Network Solution. *AIDS research and treatment*. 2012;2012:815823.  
20  
21 [20] Cislaghi B, Denny EK, Cisse´ M, Gueye P, Shrestha B, Shrestha PN, et al. Changing Social Norms: the  
22 Importance of “Organized Diffusion” for Scaling Up Community Health Promotion and Women  
23 Empowerment Interventions. *Prevention Science*. 2019;p. 1–11.  
24  
25 [21] Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden popula-  
26 tions for HIV surveillance. *Aids*. 2005;19:S67–S72.  
27  
28 [22] Smith R, Rossetto K, Peterson BL. A meta-analysis of disclosure of one’s HIV-positive status, stigma  
29 and social support. *AIDS care*. 2008;20(10):1266–1275.  
30  
31 [23] Watters JK, Biernacki P. Targeted sampling: options for the study of hidden populations. *Social*  
32 *problems*. 1989;36(4):416–430.  
33  
34 [24] Benjamin-chung J, Abedin J, Berger D, Clark A, Jimenez V, Konagaya E, et al. Spillover effects on  
35 health outcomes in low- and middle-income countries: a systematic review. *International Journal of*  
36 *Epidemiology*. 2017;(September):1–26.  
37  
38 [25] Neema S, Atuyambe LM, Otolok-Tanga B, Twijukye C, Kambugu A, Thayer L, et al. Using a clinic  
39 based creativity initiative to reduce HIV related stigma at the Infectious Diseases Institute, Mulago  
40 National Referral Hospital, Uganda. *African health sciences*. 2012;12(2):231–239.  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

- 1  
2  
3  
4 [26] Smith Fawzi MC, Siril H, Li N, McAdam K, Ainebyona D, McAdam E, et al. Agents of Change:  
5 stepped wedge randomized controlled trial of the NAMWEZA intervention with people living with  
6 HIV and members of their social networks. *BMJ Global Health*. 2019;.
- 7  
8  
9 [27] Kaaya S, Arthur G, Siril H, Lienert J, Li N, McAdam K, et al. Leadership of PLWH as Agents of  
10 Change: comparing HIV-related risk behavior of people attending ART clinics in Dar es Salaam with  
11 members of their social networks. In: 20th International AIDS Conference; 2014..
- 12  
13  
14 [28] Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC medical research*  
15 *methodology*. 2006;6(1):54.
- 16  
17  
18 [29] Zou J, Yamanaka Y, John M, Watt M, Ostermann J, Thielman N. Religion and HIV in Tanzania:  
19 influence of religious beliefs on HIV stigma, disclosure, and treatment attitudes. *BMC public health*.  
20 2009;9(1):75.
- 21  
22  
23 [30] McCarthy B. Appreciative Inquiry: an alternative to behaviour management. *Dementia*.  
24 2017;16(2):249–253.
- 25  
26  
27 [31] Siril HN, Kaaya SF, Fawzi MKS, Mtisi E, Somba M, Kilewo J, et al. CLINICAL outcomes and loss to  
28 follow-up among people living with HIV participating in the NAMWEZA intervention in Dar es  
29 Salaam, Tanzania: a prospective cohort study. *AIDS research and therapy*. 2017;14(1):18.
- 30  
31  
32 [32] Carey MP, Schroder KE. Development and psychometric evaluation of the brief HIV Knowledge  
33 Questionnaire. *AIDS education and prevention*. 2002;14(2):172–182.
- 34  
35  
36 [33] Haile BJ, Chambers JW, Garrison JL. Correlates of HIV knowledge and testing: Results of a 2003  
37 South African survey. *Journal of Black Studies*. 2007;38(2):194–208.
- 38  
39  
40 [34] Aarø LE, Breivik K, Klepp KI, Kaaya S, Onya HE, Wubs A, et al. An HIV/AIDS knowledge scale for  
41 adolescents: item response theory analyses based on data from a study in South Africa and Tanzania.  
42 *Health education research*. 2011;26(2):212–224.
- 43  
44  
45 [35] Peltzer K, Matseke G, Mzolo T, Majaja M. Determinants of knowledge of HIV status in South Africa:  
46 results from a population-based HIV survey. *BMC public health*. 2009;9(1):174.
- 47  
48  
49 [36] Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. *Pharmaceutical Statistics: The*  
50

- Journal of Applied Statistics in the Pharmaceutical Industry. 2009;8(1):50-61.
- [37] Easterly W. How the millennium development goals are unfair to Africa. *World development*. 2009;37(1):26–35.
- [38] Barbarin OA, Khomo N. Indicators of economic status and social capital in South African townships: What do they reveal about the material and social conditions in families of poor children? *Childhood*. 1997;4(2):193–222.
- [39] Smith JA, McPherson M, Smith-Lovin L. Social distance in the United States: Sex, race, religion, age, and education homophily among confidants, 1985 to 2004. *American Sociological Review*. 2014 Jun;79(3):432-56.
- [40] Lesko CR, Edwards JK, Cole SR, Moore RD, Lau B. When to censor? *American journal of epidemiology*. 2017;187(3):623–632.
- [41] Hernán MA. The hazards of hazard ratios. *Epidemiology (Cambridge, Mass.)*. 2010;21(1):13.
- [42] VanderWeele T. *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press; 2015.
- [43] Forastiere L, Mealli F, Vanderweele TJ. Identification and Estimation of Causal Mechanisms in Clustered Encouragement Designs: Disentangling Bed Nets Using Bayesian Principal Stratification Identification and Estimation of Causal Mechanisms in Clustered Encouragement Designs : Disentangling Be. *Journal of the American Statistical Association*. 2016;111(514):510–525. Available from: <http://dx.doi.org/10.1080/01621459.2015.1125788>.
- [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.
- [45] Shalizi CR, Thomas AC. Homophily and contagion are generically confounded in observational social network studies. *SMR*. 2012;40(2):211–39.
- [46] De La Haye K, Robins G, Mohr P, Wilson C. Homophily and contagion as explanations for weight similarities among adolescent friends. *Journal of Adolescent Health*. 2011;49(4):421–427.
- [47] Vanderweele TJ, Hong G, Jones SM, Brown JL. Mediation and spillover effects in group-randomized

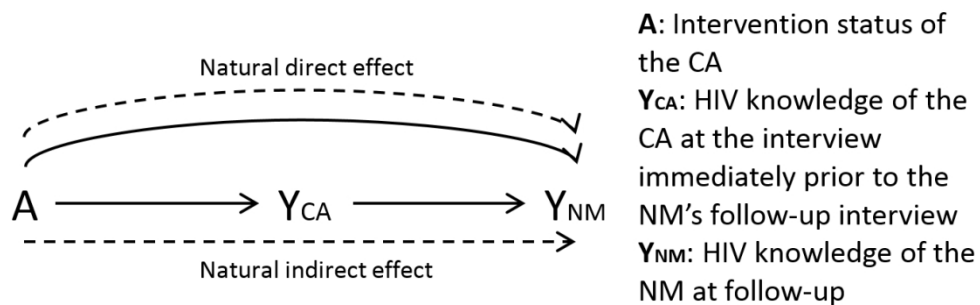
- 1  
2  
3 trials: a case study of the 4Rs educational intervention. *Journal of the American Statistical Association*.  
4 2013;108(502):469–482. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23997375>.  
5  
6  
7 [48] Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal  
8 interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological*  
9 *methods*. 2013;18(2):137.  
10  
11  
12 [49] Leung MP. Treatment and spillover effects under network interference. *Review of Economics and*  
13 *Statistics*. 2016;p. 1–42.  
14  
15  
16 [50] Latkin CA, Davey-Rothwell MA, Knowlton AR, Alexander KA, Williams CT, Boodram B. Social  
17 network approaches to recruitment, HIV prevention, medical care, and medication adherence. *Journal*  
18 *of Acquired Immune Deiciency Syndromes*. 2013;63:1–9.  
19  
20  
21 [51] Galea S, Tracy M. Participation rates in epidemiologic studies. *Annals of epidemiology*.  
22 2007;17(9):643–653.  
23  
24  
25 [52] Berkman LF. Social support, social networks, social cohesion and health. *Social work in health care*.  
26 2000 Sep 26;31(2):3-14.  
27  
28  
29 [53] Langford CPH, Bowsher J, Maloney JP, Lillis PP. Social support: a conceptual analysis. *Journal of*  
30 *advanced nursing*. 1997;25(1):95–100.  
31  
32  
33 [54] Schneider JA, Zhou AN, Laumann EO. A new HIV prevention network approach: sociometric peer  
34 change agent selection. *Social science & medicine*. 2015;125:192–202.  
35  
36  
37 [55] He P, Eriksson F, Scheike TH, Zhang MJ. A proportional hazards regression model for the  
38 subdistribution with covariates-adjusted censoring weight for competing risks data. *Scandinavian*  
39 *Journal of Statistics*. 2016 Mar;43(1):103-22.  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

## 11. Figure Captions

Fig 1. Schematic of natural direct effect (NDE) and natural indirect effect (NIE). The NDE indicates the increase in NMs' HIV knowledge happens without a concomitant increase in their CA's HIV knowledge. The NIE, on the other hand, indicates that the increase in NMs' HIV knowledge is mediated by their CA's HIV knowledge increasing. Solid lines indicate paths of causality between variables. Dashed lines represent the line or lines composing the effect of interest.

For peer review only

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60





## STROBE (Strengthening The Reporting of OBServational Studies in Epidemiology) Checklist

A checklist of items that should be included in reports of observational studies. You must report the page number in your manuscript where you consider each of the items listed in this checklist. If you have not included this information, either revise your manuscript accordingly before submitting or note N/A.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).

Section and Item	Item No.	Recommendation	Reported on Page No.
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	
<b>Introduction</b>			
Background/Rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
<b>Methods</b>			
Study Design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed  <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	

Section and Item	Item No.	Recommendation	Reported on Page No.
Data Sources/ Measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study Size	10	Explain how the study size was arrived at	
Quantitative Variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	
Statistical Methods	12	(a) Describe all statistical methods, including those used to control for confounding	
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive Data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome Data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	

Section and Item	Item No.	Recommendation	Reported on Page No.
Main Results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other Analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key Results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
<b>Other Information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Once you have completed this checklist, please save a copy and upload it as part of your submission. DO NOT include this checklist as part of the main manuscript document. It must be uploaded as a separate file.**

# BMJ Open

## Evaluating spillover of HIV knowledge from study participants to their network members in a stepped-wedge behavioral intervention in Tanzania

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-033759.R4
Article Type:	Original research
Date Submitted by the Author:	13-Jul-2020
Complete List of Authors:	Lienert, Jeffrey; University of Pennsylvania Perelman School of Medicine, Center for Health Care Innovation; Philadelphia VAMC, CHERP Fawzi, Mary C; Harvard Medical School, Global Health and Social Medicine McAdam, Keith; London School of Hygiene and Tropical Medicine, Epidemiology Kaaya, Sylvia; Muhimbili University of Health and Allied Sciences, School of Medicine Liu, Yuanyuan Todd, Jim; London School of Hygiene and Tropical Medicine, Population Health Andrew, Irene; Management and Development for Health, Onnela, Jukka Pekka; Harvard T.H. Chan School of Public Health, Biostatistics
<b>Primary Subject Heading</b>:	HIV/AIDS
Secondary Subject Heading:	Global health, Public health
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, STATISTICS & RESEARCH METHODS

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3  
4  
5 Evaluating spillover of HIV knowledge from study participants to their  
6 network members in a stepped-wedge behavioral intervention in Tanzania  
7  
8

9 Jeffrey Lienert<sup>a,b,j</sup>, Mary C. Smith Fawzi<sup>c</sup>, Keith McAdam<sup>d</sup>, Sylvia Kaaya<sup>e</sup>, Yuanyuan Liu<sup>i</sup>, Jim Todd<sup>g</sup>,  
10 Irene Andrew Lema<sup>f</sup>, Jukka-Pekka Onnela<sup>h</sup>

11  
12 <sup>a</sup>Perelman School of Medicine, University of Pennsylvania <sup>b</sup>Center for Health  
13 Equity Research and Promotion, Philadelphia VA <sup>c</sup>Department of Global Health  
14 and Social Medicine, Harvard Medical School <sup>d</sup>Department of Clinical Tropical  
15 Medicine, London School of Hygiene and Tropical Medicine <sup>e</sup>Department of Psychiatry and  
16 Mental Health, Muhimbili University of Health and Allied Sciences <sup>f</sup>Management and  
17 Development for Health <sup>g</sup>Department of Population Health, London School of  
18 Hygiene and Tropical Medicine <sup>h</sup>Department of Biostatistics, Harvard T. H. Chan School of  
19 Public Health <sup>i</sup>University of Texas School of Public Health  
20 <sup>j</sup>Corresponding Author: Jeffrey Lienert, jeffrey.lienert@penntmedicine.upenn.edu  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

52 Preprint submitted to BMJ Open

8 July, 2020

## Abstract

**Objectives:** We aim to describe the social network members of participants of a behavioral intervention, and examine how the effects of the intervention may spillover among network members.

**Design:** Secondary analysis of a step-wedge randomized controlled trial.

**Setting:** Change Agents (CAs) were recruited from waiting rooms of HIV treatment facilities in Dar es Salaam, Tanzania, and their Network Members (NMs) were recruited directly by CAs.

**Participants:** We enrolled 662 CAs in an HIV behavioral intervention. They, along with 710 of their NMs, completed baseline and follow-up interviews from 2011 to 2013.

**Primary and Secondary Outcomes:** The primary outcome of this study was change in NMs' HIV knowledge, and the secondary outcome was whether the NM was lost to follow-up.

**Results:** At baseline, many characteristics were different between NMs and CAs. We found a number of NM characteristics significantly associated with follow-up of NMs, particularly female gender (OR=1.64, 95% CI: 1.02,2.63) and HIV knowledge (OR=20.0, 95% CI: 3.70,125); only one CA variable was significantly associated with NM follow-up: having a private source of water (OR=2.17, 95% CI: 1.33,3.57). The 14.2% increase in NMs' HIV knowledge was largely due to CAs feeling empowered to pass on prior knowledge, rather than transmitting new knowledge to their NMs.

**Conclusions:** Characteristics of social network members of persons living with HIV PLH may play a role in study retention. Additionally, the HIV knowledge of these NMs increased largely as a function of CA participation in the intervention, suggesting that intervening among highly-connected individuals may maximize benefits to the potential population for whom spillover can occur.

*Keywords:*

mediation, natural direct effect, natural indirect effect, follow-up, retention, hidden population

## 1. Strengths and Limitations of This Study

- The recruitment method, focused on social network members of persons living with HIV, allowed us to access an at-risk population typically not easily-accessible via other means, such as direct approach by the study team rather than someone known to the participant.
- The study design of the trial in which this study is nested was an ideal situation in which to employ a mediation analysis that informs our understanding of how this and similar intervention effects may spillover in a population.
- The greater-than-ideal dropout rate of NMs was both a strength in that it allowed us to examine factors associated with dropout, but also a limitation, in that the potential of differential dropout by unmeasured factors may have biased some of our results.

## 2. Introduction

Despite recent decreases in mortality, HIV/AIDS is currently the ninth leading cause of years of life lost (YLL) globally [1]. Research into eliminating HIV has focused on two fronts: biological and behavioral, which combined have decreased the years of life lost due to HIV by an estimated 51% [1]. Behavioral interventions seek to curb transmission of HIV through reduction of risk behavior including safer sexual practices and reducing injection drug use [2, 3]. Behavioral change also impacts the effectiveness of the biological methods, given that lower adherence is known to often compromise their effectiveness [4]. However, neither biological nor behavioral methods are fully effective without the knowledge of these treatments and preventative behaviors. It is therefore important to increase HIV-related knowledge as a necessary, but not sufficient, step towards achieving the 90-90-90 treatment goals (90% diagnosis, 90% antiretroviral therapy, and 90% viral suppression among the treated) and ultimately eliminate HIV [5].

The importance of adequate HIV knowledge has been recognized for a period of time, as it is often a prerequisite for behavioral change [6]. It should be noted that increased knowledge does not necessarily translate into behavior change, as other factors such as motivation and skills also play a role [6]. Despite this, researchers conducted trials among Persons Living with HIV (PLH) which led to increased HIV knowledge [7, 8]. These interventions were group discussions and teaching led by nurses, and motivational



1  
2  
3 interviewing, respectively. Researchers also found that the intervention reduced HIV-risk behaviors  
4 concomitantly with an increase in HIV knowledge [8].  
5

6  
7 1. Knowledge gained by participants in these trials can also be freely shared with members of their  
8 social network, in what is known as a spillover effect [9]. Specifically, a spillover effect (also known as an  
9 indirect or disseminated effect) is one person's exposure affecting another's outcome [10]. In the context of  
10 a social network spillover intervention, it corresponds to an individual who was unexposed to an intervention  
11 changing their behavior because they were socially-connected to an individual who did receive the  
12 intervention. This is distinct from what is sometimes called behavioral spillover, where changes in a person's  
13 behavior affect other behaviors of that same person [11]. For injection drug users, HIV prevention  
14 educational interventions were demonstrated to have spillover effects of HIV prevention education, and  
15 subsequent reduced rates of risky behaviors [12]. Studies have also used proxy variables for social network  
16 ties such as inviting social network members to watch educational programming [13] or time spent shopping  
17 at the market [14] to evaluate spillover effects for HIV knowledge, generally finding evidence for spillover.  
18 However, spillover in HIV knowledge between known social network ties generally remains understudied,  
19 particularly in sub-Saharan Africa [15]. We therefore aim to determine whether social network members  
20 those receiving an HIV behavioral/knowledge intervention also increase their HIV knowledge.  
21  
22  
23  
24  
25  
26  
27  
28  
29

30 New knowledge can come from a variety of sources, one of the most important of which is a person's  
31 social network [16, 17]. Social networks are of particular import because new knowledge can lead to  
32 cascades of behavior change, where people subsequently educate those in their social network, in what is  
33 known as social influence [18, 19, 20]. This has been examined in participant-driven interventions, where  
34 initially-recruited participants educate members of their social network one-on-one [21]. Characteristics  
35 such as knowing the HIV status of network members has been shown to be the most important predictor of  
36 engaging in prevention advocacy [22]. Work on diffusion through social networks, how a belief or behavior  
37 can be "contagious" within a network, has shown that spreading intervention effects beyond the initial study  
38 population can improve the cost-effectiveness of these interventions [23]. These findings imply that certain  
39 aspects of knowledge or behavior may spread more or less efficiently through networks comprising  
40 individuals with specific characteristics, which may need to be accounted for in network interventions. For  
41 instance, networks comprising many at-risk individuals who are HIV-negative may not be as receptive to  
42 change in behaviors as networks comprising a mix of those with and without HIV. Additionally, the  
43 networks of PLH are often difficult networks to ascertain [24], due to the continued stigma of HIV and AIDS  
44  
45  
46  
47  
48  
49  
50  
51  
52

1  
2  
3 in many settings [25]. Because of this, if a PLH or person at-risk of acquiring HIV does not want to  
4 participate in an investigator-initiated intervention, there is little recourse other than information transmitted  
5 via social networks, or targeted sampling techniques which are not always effective (e.g. Respondent Driven  
6 Sampling) [26]. This is particularly important in low- and middle-income countries, as a recent systematic  
7 review found only 54 studies researching spillover effects in these settings (out of approximately 750) [27].  
8 Therefore, understanding exactly how information spreads from participants in an intervention to members  
9 of their social network, who may be largely inaccessible via other means, is important for reaching the  
10 greatest number of people about HIV prevention. Understanding what makes these persons different from  
11 those who receive the intervention themselves is important, as it may point to ways in which to increase  
12 enrollment of these populations.  
13

14  
15 Based on the above gaps in the literature, we conduct a study on network members of PLH enrolled in a  
16 behavior change intervention [28]. The trial recruited PLH to serve as Change Agents (CAs) and to reach  
17 out to their social network members (NMs) about knowledge of HIV and safer sexual practices [29]. Our  
18 goals in this study were threefold: 1) to understand how the NMs differed from their CAs, 2) understand  
19 correlates of dropout for the NMs, and 3) to understand by what method HIV knowledge transferred to the  
20 NMs from the CAs. Understanding how the information and behaviors are shared within social networks  
21 will allow HIV researchers and others to take advantage of this knowledge and improve upon prevention  
22 interventions in the future.  
23  
24

### 25 3. Methods

#### 26 3.1. Study Population

27  
28 We analyze social network data from the Agents of Change trial [30], which was a stepped-wedge  
29 randomized controlled trial [31] that enrolled PLH to become Change Agents (CAs) by informing members  
30 of their social network (NMs) about knowledge of HIV and safer sexual practices. Here, we define CAs  
31 based on the *potential* for PLH to become so - by self-selecting into the study, PLH identify themselves as  
32 potential CAs. Although we refer to them as 'CAs' throughout, participants in the trial enrolled with varying  
33 levels of ability to act as a Change Agent. Through receiving the *NAMWEZA* intervention, we hypothesize  
34 that CAs will be able to truly self-actualize and subsequently act as Change Agents in their community.  
35

36  
37 CAs were recruited from the waiting rooms of HIV care and treatment centers in Dar es Salaam, Tanza-  
38 nia, and we received written consent from each CA. Participants completed a baseline questionnaire and  
39

1  
2  
3 were randomized to one of three waves in which to receive the intervention. At baseline, participants were  
4 also asked to recruit up to three members of their social networks who they felt were at particularly high risk  
5 of contracting or spreading HIV. We obtained written consent from these nominated network members (NMs).  
6 NMs could be either HIV positive or negative, and they were given a baseline survey. The NM was only  
7 aware the CA was a participant in the intervention if the CA shared this information with them, which  
8 many did not due to HIV-related stigma [32]. Each CA therefore formed a CA-NM dyad with each NM they  
9 recruited, and if they recruited more than one NM, formed a set of CA-NM dyads with a common CA.

10  
11 As fits a stepped-wedge RCT, all CAs eventually received the intervention, but were randomized to  
12 *when* they received it. These waves each lasted 12 weeks, at which point the next wave began and another  
13 group of CAs received the intervention. Within each wave, the intervention comprised 10 weekly structured  
14 sessions aimed at empowering PLH to become HIV prevention change agents in their communities. The  
15 sessions aimed to increase psychosocial and communication skills through an Appreciative Inquiry approach  
16 [33]. Within one month of each wave of the intervention, CAs were given follow-up surveys. Across all  
17 waves, the interventions lasted from November 2010 to January 2014, and the final interviews were  
18 conducted in March 2014. For more information on the study design, we direct interested readers to Smith-  
19 Fawzi et al., 2019 [29].

20  
21 The NMs did not receive any intervention at any point during the study. Rather, their intervention status  
22 flowed from the intervention status of their CAs. Due to operational difficulties, each NM was surveyed two  
23 times during the study: baseline and after the first wave, rather than baseline and one after each wave. In  
24 this way, all demographic and contextual variables were measured at baseline. We use this interview of NMs  
25 as the division between “exposed” and “unexposed”. At the time of an NM’s follow-up interview, their  
26 respective CA may or may not have undergone the intervention. Therefore, the NMs were divided into  
27 “exposed” (N=381) and “unexposed” (N=329) groups based on whether their respective CA had completed  
28 their intervention at the time of the NM’s follow-up interview.

29  
30 Inclusion criteria for the follow-up analyses were for CAs who completed at least one follow-up inter-  
31 view and had at least one NM who completed at least one follow-up questionnaire. In this way, outcomes  
32 could be computed for the CA and NM of each dyad. NMs approached by a CA, but who did not participate  
33 in the study were not recorded since it was not feasible to obtain this information from study participants  
34 themselves. As we lost some CAs and NMs to follow up, we completed our analyses without their data,  
35 assuming it to be Missing Completely at Random (MCAR). During this study, there was little loss-to-follow-  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

1  
2  
3 up among the CAs (< 10%), but much higher loss among the NMs (36.8%) [34]. Given an NM or CA was  
4 not lost to follow-up, complete information was available on all additional variables, including exposure,  
5 outcome, and covariates. In sum, our sample comprises 662 CAs and 710 NMs, meaning each CA recruited  
6  
7 1.07 NMs on average out of a possible 3, and 44 CAs nominated at least two NMs.  
8  
9

### 10 11 *13.2.HIV knowledge*

12  
13 To assess HIV knowledge of CAs and NMs, we used the brief HIV knowledge questionnaire [35]. This  
14 scale comprises 18 questions, which focus on an individual's knowledge of how HIV can spread and other  
15 characteristics of the virus and AIDS. This knowledge is crucial for an individual's subsequent safe-sex  
16 practices to reduce the risk of transmission. The original population comprised three different groups: two  
17 groups of low-income women, and one of women and men receiving psychiatric treatment. In these popula-  
18 tions, questions on the measure had a Chronbach's alpha of 0.78. This instrument has been used previously  
19 in sub-Saharan Africa and has demonstrated reasonable reliability of 0.75 Cronbach's alpha in South Africa  
20 among a convenience sample of 429 members of the African Methodist Episcopal church [36]. It has also  
21 been translated to Swahili, with only minor differential item functioning [37]. This indicates that the  
22 measure performs adequately in other, similar populations. In the present study the Cronbach's alpha is  
23 0.71. Because the average baseline score of HIV knowledge was 0.80, with a preponderance of scores of  
24 1.00, the main indicator of knowledge was not normally distributed, and therefore a continuous predictor  
25 was not ideal [38,39]. We therefore summarize this measure as "Complete HIV Knowledge", a dichotomous  
26 variable for whether the participant correctly answered all questions.  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36

### 37 *13.3.Demographic and contextual variables*

38  
39 In addition to HIV knowledge, we are interested in a set of demographic and contextual variables that  
40 may help explain some of the trends observed. In terms of demographic variables, we include age, sex,  
41 employment status, marital status, and self-identified HIV status. We also include education, which we di-  
42 chotomize at 7 years or more - a level commensurate with elementary education. We selected this cutoff  
43 because it coincides with the millennium development goal (MDG) of increasing primary education com-  
44 pletion [40].  
45  
46  
47  
48  
49

50 Contextually, we included two additional variables: having a private source of water, and the number  
51 of persons sleeping in the participant's home. The first was also based on a MDG, and indicates participants  
52  
53

1  
2  
3 with access to safe drinking water [40]. This is a proxy for the economic security of the participant. The  
4 number of persons sleeping in the participant's home is also a non-monetary indicator of their material and  
5 social resources [41]. These variables combined give a more thorough picture of the participant's economic  
6 status than employment alone.  
7  
8  
9

#### 10 *13.4. Statistical Analysis*

11  
12 The first analysis was to assess homophily, defined as the extent to which CAs and NMs were similar to  
13 one another in terms of a number of sociodemographic characteristics and HIV-related risk factors. Because  
14 CAs and NMs self-selected into their respective group, and only CAs were directly randomized (with their  
15 NMs being randomized along with them), differences between the groups were to be expected. However,  
16 we only examined homophily of baseline characteristics rather than of outcome, because comparison of  
17 outcomes between CAs and NMs would remove the benefits of randomization. Because CAs and NMs self-  
18 selected into the study and were not randomized to CA/NM status, we do not *a priori* expect them to be  
19 completely similar. In addition, multiple NMs could share a CA and would therefore not be independent due  
20 to the shared variation and latent characteristics of having the same CA. We therefore assessed statistical  
21 significance of homophily on the set of CA-NM dyads using a permutation test, a non-parametric test which  
22 has no distributional assumptions. For continuous variables, the difference between the CA and NM was  
23 calculated, and for categorical variables, whether the CA and NM were concordant or discordant was  
24 recorded. We then randomly permuted CA-NM ties (keeping number of ties per CA constant), and then  
25 recalculated the difference or percent concordant, respectively, 1,000 times. We then examine the percentile  
26 of the observed difference relative to the permuted differences [42]. Analyses were run using R v3.1.1.  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36

37 To accomplish our second aim of understanding what was associated with NMs completing their follow-  
38 up interview, we fit a log-binomial regression to determine predictors of follow-up. In this regression, we  
39 use all the variables listed above, as well as whether the CA remained in the study for its full duration. In  
40 order to examine the association between these same variables and the time of follow-up, we also employed  
41 a Cox regression. In the Cox regression, start of follow-up was defined as the time at which the NM  
42 completed their baseline interview. The outcome here was whether the NM completed a follow-up interview.  
43 NMs who were lost to follow-up were censored at the time of their latest interview [43,44].  
44  
45  
46  
47  
48

49 Finally, since the trial showed beneficial effects on the HIV knowledge of the NMs following the inter-  
50 vention [30], we aim to elucidate exactly what caused the HIV knowledge of NMs to increase - either CAs  
51  
52

1  
2  
3 gaining knowledge through the intervention and sharing it, or the CAs being empowered by the intervention  
4 to share existing knowledge. As the wedge in which the CA received the *NAMWEZA* intervention was  
5 randomized, we treat each NM as being randomized to exposure to *NAMWEZA* at the same time as their  
6 CA. This randomization scheme allows us to explore the spillover effect of CAs' HIV knowledge onto their  
7 CA. This randomization scheme allows us to explore the spillover effect of CAs' HIV knowledge onto their  
8 respective NMs via a mediation analysis. These pathways represent different types of spillover effects: the  
9 exposure or outcome of one person affecting the outcome of another person.

13 As shown by VanderWeele et al (2015), social network spillover effects in the case of dyadic  
14 relationships can be broken down into concepts from mediation analysis: direct and indirect effects (Figure  
15 1) [45]. This method has since been used for novel evaluations of spillover effects [46,47,48]. Although  
16 previous studies showed that this same type of analysis cannot be done on full network data, the data in this  
17 study consisted of only dyads, the CAs and their NMs, so in this case the method is appropriate [49,50].

22 The method parses social influence into direct and indirect effects. The natural indirect effect (NIE) is  
23 the effect by which the intervention changes the CA's HIV knowledge, which then changes the NM's HIV  
24 knowledge. The NIE is similar to the effect observed in many participant-driven interventions: an initial  
25 participant receives the intervention, increasing their knowledge, and they subsequently pass their increased  
26 knowledge to members of their social network [21]. The natural direct effect (NDE) is the effect of receiving  
27 the intervention has on an NMs outcome, irrespective of the CA's outcome (in this case HIV knowledge).  
28 For instance, this could occur if CAs begin with good knowledge of HIV, and the intervention empowers  
29 them to convey knowledge they already had to their NMs. Although the intervention does not increase their  
30 HIV knowledge, it is still useful to the CAs, as it empowers them to act as CAs in their community. In order  
31 to estimate these effects, the published SAS macro developed by VanderWeele and colleagues was used,  
32 and analyses were run using SAS v9.2 (SAS Institute, Cary, NC) [51,52].

41 Importantly, this analysis requires a number of assumptions and applies to dyads only when these  
42 assumptions are met. One assumption of this analysis is that the dyads are independent, which is violated  
43 here; if a CA recruited more than one NM, the multiple CA-NM dyads involving the same CA would not  
44 be independent. To address this, we performed the analysis after randomly removing NMs until each CA  
45 had only a single NM. This resulted in removing 48 NMs, just 6.7% of the population. We found that the  
46 point estimates were nearly identical, but that the confidence intervals were slightly larger due to the reduced  
47 sample size. No coefficients changed from significant to non-significant in this analysis (data not shown).  
48  
49  
50

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
A second, related assumption is that of *partial interference*, that the effects in one cluster does not affect another cluster - here, one CA-NM dyad affecting another [53]. This could occur if two NMs of different CAs happen to know one-another outside of the study, one has a CA who was randomized to an earlier wedge, and shares what they know of it with the otherwise-unexposed NM. However, due to the large size of Dar es Salaam, and the number of HIV treatment clinics in which recruitment occurred, we expect few CAs or NMs to know one another outside of the study (other than NMs knowing the CA who recruited them), limiting the potential for partial interference.

15  
16  
17  
18  
19  
20  
21  
22  
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 any confounding variables [45,46]. 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 the variables used in the log-binomial regression.

23  
24  
25  
26  
27  
28  
29  
30  
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.

### 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 *13.5. Patient and Public Involvement*

Patients were not involved in the design of this study. As part of the recruitment process, patients were instructed to recruit members of their social networks who they felt were at high risk of HIV infection - in this way participants could guide the recruitment of NMs to those most at-risk. We currently have no plans to disseminate the results of this study to participants.

## 4. Results

The *NAMWEZA* study recruited 662 Change agents (CA), of whom 453 (68.4%) completed at least one follow-up questionnaire. These 662 CAs in turn recruited 710 network members (NM) who took the baseline questionnaire. Of the 710 NMs, 449 (63.2%) also completed a follow-up questionnaire (Table 1). At baseline, CAs were on average older than their NMs. More of the NMs were employed than their CAs (69.3% vs. 55.0%,  $p < 0.001$ ), but were less likely to have at least 7 years of education (52.0% vs. 52.3%,  $p < 0.001$ ). Only

12.3% of NMs were HIV-positive, compared to *all* CAs ( $p < 0.001$ ). Complete data was obtained at baseline for all CAs and NMs.

Characteristic	Number of NMs (%) or Mean (SD) (N=710)	Number of CAs (%) or Mean (SD) (N=662)	P-value
Age	33.0 (11.1)	38.9 (9.7)	<0.001
Female	380 (53.7%)	349 (53.9%)	0.89
Employed	490 (69.3%)	356 (55.0%)	<0.001
At least 7 years education	369 (52.0%)	584 (82.3%)	<0.001
Complete HIV knowledge	638 (89.9%)	598 (90.4%)	0.65
Persons sleeping in home	5.00 (5.77)	3.87 (3.72)	<0.001
Married	373 (52.7%)	338 (51.1%)	0.56
HIV Positive	87 (12.3%)	662 (100%)	N/A
Private source of water	309 (43.7%)	263 (39.7%)	0.19

Table 1: Demographic characteristics at baseline, with the results of a permutation test for homophily. Specifically, for each CA- NM dyad, either a difference (for continuous variables) or concordance (for dichotomous variables) is calculated. For example, if a CA was 39 years old, and their NM was 25 years old, the difference would be 14 years old. If a CA was male and their NM was Female, the pair would be discordant for sex. CA-NM pairs were then randomly reshuffled, the edge-wise characteristics recalculated, and the observed difference compared to the distribution of randomized differences.

Risk ratios (RRs) obtained via log-binomial regression indicated that characteristics of both the CAs and the NMs significantly predicted loss to follow-up (Table 2). The NM being female (RR=1.44, 95% CI: 1.05,1.97), having complete HIV knowledge (RR=10, 95% CI: 2.33,42), being employed (RR=1.43, 95% CI: 1.08,1.89), and being married (RR=1.55, 95% CI: 1.03,2.33) were all significant predictors of increased odds of completing a follow-up interview. Each additional person sleeping in the home of the NM per room used for sleeping reduced the odds of follow-up (RR=0.85, 95% CI: 0.74,0.98) as did the NM living with HIV (RR=0.40, 95% CI: 0.17,0.96). CA having a private water source was significantly associated with increased odds of the NM being followed-up (RR=2.07, 95% CI: 1.25,3.42), even after controlling for the NM having a private source of water. This was the only CA-specific variable that significantly predicted an NM's follow-up.

The Cox proportional hazard model showed similar results to the logistic regression, but with fewer significant results. Only the NM being married (HR=1.28, 95% CI: 1.01,1.64) and the NM living with HIV (HR=0.71, 95% CI: 0.51,0.99) significantly predicted the time-to-follow-up of the NMs. Additionally, all of the hazard ratios were closer to the null relative to the corresponding odds ratios.



We assess the mechanism of spillover effects on HIV knowledge, the main effect of the intervention on the NMs. The analysis showed that NMs who were exposed to intervention via their CA had an increase in HIV knowledge of 12.9% on average (95% CI: 0.06, 0.20) from a baseline percentage of 78% [26]. This was broken down into a natural indirect effect (NIE) of 0.6% (95% CI: -0.06%,2.0%), which is the effect the intervention had via the CA's HIV knowledge, and a natural direct effect (NDE) of 12.3% (95% CI: 6.1%,19.3%), which is the effect the CA's participation had on the NM's HIV knowledge, irrespective of the CA's HIV knowledge. In other words, of the 12.9% increase in NM's having complete HIV knowledge, 12.3% occurred without a concomitant increase in their CA's HIV knowledge. In other words, their HIV knowledge increase was not mediated by the increase in HIV knowledge of their CA. This did not change when we used only one NM per CA.

## 5. Discussion

Based on results from the NAMWEZA trial in Tanzania, we have shown a number of novel findings regarding the network members of HIV intervention participants, correlates of their retention in the study, and the evidence for spillover of HIV knowledge from CAs to NMs. These findings can inform the design of interventions in the future to maximally enroll and retain participants and ensure that intervention information is transmitted from the study participants to members of their social networks.

Characteristic (N=459)	Adjusted RR (95% CI)	Adjusted HR (95% CI)
Characteristics of NMs		
Gender (Female)	1.44* (1.05,1.97)	1.18 (0.94,1.50)
Difference in age of NM and CA (per year)	1.01 (0.99,1.03)	1.01 (0.995,1.02)
Age of NM	0.98 (0.95,1.02)	0.99 (0.97,1.04)
Complete HIV knowledge (vs. none)	10* (2.33,42)	2.20 (0.97,5.01)
Employed	1.43* (1.08,1.89)	1.15 (0.89,1.50)
Each additional person sleeping in home per room used for sleeping	0.85* (0.74,0.98)	0.92 (0.83,1.01)
Married	1.55* (1.03,2.33)	1.28* (1.01,1.64)
Living with HIV	0.40* (0.17,0.96)	0.71* (0.51,0.99)
Having a private source of water	0.89 (0.59,1.34)	0.97 (0.77,1.22)
Characteristics of CAs		
Gender (Female)	1.27 (0.76,2.08)	1.07 (0.84,1.37)
Complete HIV knowledge (vs. none)	0.36 (0.07,2.04)	0.64 (0.29,1.43)
Having a private source of water	2.07* (1.25,3.42)	0.97 (0.77,1.22)
Being employed	1.54 (0.95,2.50)	1.14 (0.90,1.43)
Being married	1.25 (0.77,2.04)	1.11 (0.88,1.43)
CA Lost to follow-up	1.06 (0.84,1.33)	1.02 (0.72,1.43)

Table 2: Results of multivariate log-binomial regression and Cox-proportional hazard models on the dichotomous outcomes of whether the NM completed a follow-up questionnaire, and the continuous outcome of time-to-completion of follow-up questionnaire, respectively. “\*” indicates significant at the  $p < 0.05$  level.

1  
2  
3  
4  
5 We found several significant differences between CAs and NMs they recruited. On average, CAs  
6 were older, less likely to be employed, and more educated. All of this suggests that by having CAs recruit  
7 from their social network, we were able to recruit a set of social network members different in multiple ways  
8 from those found in the waiting rooms in HIV treatment clinics (unsurprising, given that many NMs were  
9 not HIV-positive) [9]. This may in part be due to the fact that when recruiting NMs, CAs were instructed to  
10 recruit those who may be at particular risk of contracting or spreading HIV. Therefore, CAs likely did not  
11 pick random members of their social network, but those who were at high risk. This finding corroborates  
12 the claim by Latkin et al. (2013) that sampling via enumeration of network members by initial participants  
13 recruits a more diverse sample, as our sample of NMs was not composed of only those who were HIV-  
14 positive, but also many who were HIV-negative [54]. Additionally, the lower likelihood of employment of  
15 the CAs indicates they may have had more time for participating in the intervention [55]. This means that  
16 future studies may need to tailor their interventions to work with the schedules of employed persons to  
17 increase participation.  
18  
19  
20  
21  
22  
23  
24  
25

26 Although the study design potentially accessed a separate slice of the population than other methods  
27 vis-a'-vis the population of those at risk for contracting HIV, it may have been at least partially responsible  
28 for the large loss of follow-up of the NMs. It is important to note that more NM characteristics than CA  
29 characteristics were significantly predictive of follow-up in both the logistic and Cox proportional hazards  
30 models. The one CA characteristic which did predict NMs completing a follow-up interview, having a  
31 private source of water, was a proxy for CA's socioeconomic status. This may have been because CAs with  
32 greater resources may have had more time available to pass information to their NMs, retaining the NM's  
33 interest longer [56]. We also note that less HIV knowledge and living with HIV predicted reduced likelihood  
34 of completing a follow-up questionnaire among NMs, which means that those who might have benefited  
35 most from spillover of the intervention were more likely to discontinue their involvement. This does not  
36 mean that they did not receive any spillover, only that it was not recorded. Our estimate of the magnitude of  
37 the spillover may therefore be biased towards the null. This is also problematic more generally for  
38 interventions of this nature as the very people the intervention aims to benefit may not stay with the program.  
39  
40  
41  
42  
43  
44  
45  
46  
47

48 The HIV knowledge gain experienced by the NMs was largely due to the NDE; i.e. knowledge spilled-  
49 over as a result of the CAs participating in NAMWEZA, not because the CAs' HIV knowledge increased,  
50 and they passed this new knowledge to their NM(s). This is most likely because CAs had a high average  
51  
52

1  
2  
3 HIV knowledge score at baseline (80%), so there was less room for improvements in knowledge. Following  
4 the intervention, CAs' likelihood of complete HIV knowledge did not significantly increase. Therefore, what  
5 prompted the increased knowledge of the NMs was the CA becoming empowered through the intervention  
6 to pass on their existing knowledge to their NM [57].  
7  
8  
9

10 This finding is important for future interventions: spillover effects of this intervention will likely carry  
11 over only to those directly-connected to the CAs (opposed to spreading indefinitely in a snowball effect), as  
12 the most important thing in transmitting HIV knowledge is receiving the intervention, which only CAs  
13 receive. Because NMs do not actually receive the intervention, it is unlikely that they would feel empowered  
14 to become CAs themselves (particularly because they will not receive the *NAMWEZA* sessions), thus limiting  
15 the spread of the intervention. For CAs to increase the HIV knowledge of their NMs, they would only need  
16 to become empowered to share their information. However, for the NMs to increase the HIV knowledge of  
17 their own NMs, they would need to both increase their HIV knowledge, and become empowered to share it.  
18 This is less likely to happen than just becoming empowered, and so it is unlikely, though not impossible, that  
19 this effect would continue to spread in the population. This may give insight into how to design interventions  
20 in the future; if one wants to maximize the number of people who benefit, choosing CAs who form many  
21 bridging ties in the community would maximize the potential number of links by which spillover can occur  
22 [58]. Alternatively, interventions can be designed to be self-propagating; if CAs are empowered to deliver  
23 the intervention to others, changing their own NMs into future CAs, those new NMs-turned-CAs could then  
24 deliver the intervention to a second set of NMs, again empowering them to become CAs.  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34

35 Our study is not without limitations. First, there was considerable loss to follow-up of the NMs. Al-  
36 though this actually informed our analysis of the correlates of loss to follow-up, it meant that our analysis  
37 of spillover effects may be biased. Even though the exposure was randomized, the loss to follow-up can  
38 result in selection bias if the NMs who left the study were systematically different from those who remained.  
39 As we show, the NMs who dropped out were those who would have benefited the most from the intervention  
40 because they were more vulnerable and at higher risk than those who did not drop out. Because they are  
41 likely to have greater benefits from any spillover effects, we would expect this non-random loss to follow-up  
42 to result in an underestimate of the impact of the intervention in the network members. Adjusting for  
43 censoring weights may ameliorate this issue [59]. Second, our use of Hazard Ratios (HRs) has important  
44 limitations: they are subject to selection bias, are sensitive to study period, and only provide one estimate  
45 during the study [44]. Any of these limitations could affect this analysis, hence our use of logistic regression  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 as a primary analysis. However, they remain useful as a sensitivity analysis. Third, our data did not perfectly  
4 fit the requirements of the causal mediation analysis: they were not entirely independent, since multiple NM-  
5 CA dyads shared a CA. However, when we randomly removed dyads until there were no repeated CAs, the  
6 results were qualitatively very similar, indicating that lack of independence did not unduly affect our results.  
7  
8 Fourth, although we were able to tease apart the direct and indirect effects, we are unable to determine the  
9 mechanism of the natural direct effect; the data do not allow us to specify whether NMs increased knowledge  
10 through speaking to their knowledgeable CA, through researching HIV on their own, or some other  
11 mechanism. Future work will have to be done to examine these different pathways.  
12  
13  
14  
15  
16  
17

## 18 **6. Conclusions**

19  
20 These results have implications for the potential scale-up of the *NAMWEZA* intervention, as well as  
21 future studies and interventions that focus on behavioral interventions in social networks. First, our findings  
22 of minimal similarity between CAs and their NMs indicate that this recruitment method allows us to enroll  
23 participants from portions of the population that are not represented by the CAs alone. Coupling this with a  
24 deeper understanding of the mental heuristics CAs used to select NMs (e.g. did CAs mentally search their  
25 close or peripheral network for those at-risk of HIV), may lead to different strategies for recruitment and  
26 retention, leading to stronger effects of behavioral interventions. The mediation analysis presents a  
27 compelling picture of how best to ensure the benefits of interventions reach as many people beyond the  
28 study participants as possible. Participation of CAs in the intervention resulted in positive effects on their  
29 immediate network members' HIV knowledge regardless of how the CAs responded to the intervention.  
30 While improvement in HIV knowledge may not necessarily translate to increased safe sex behavior, it can  
31 be seen as a rate-limiting step towards reducing HIV transmission since adequate knowledge is usually  
32 necessary for reducing risk behavior. Future work should examine the exact mechanisms of spillover  
33 discussed here, as that is an important clarification that could benefit future studies. Specifically, a similar  
34 setup to the study here combined with semi-structured interviews with CAs and NMs about their interactions  
35 with one another would help elucidate exactly how NMs increased their HIV knowledge (this population  
36 would likely need to comprise CAs who have disclosed their HIV status in order to prevent accidental  
37 disclosure). Future work should also examine whether increases in HIV knowledge translate to changes in  
38 behaviors which may increase one's risk of contracting HIV, particularly in Sub-Saharan Africa. The results  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

presented herein may inform approaches for increasing participation and potentially conferring greater benefits related to spillover effects in future HIV behavioral interventions.

## 7. Acknowledgements

The authors would like to thank the study participants and staff for their time, hard work, and dedication. We would also like to thank Guy Harling for their comments on the manuscript.

## 8. Contributorship

JL, MCSF, and JPO designed the analysis plan and drafted the initial paper. JPL, MCSF, KM, SK, IAL and JT conducted the underlying randomized trial, and designed it to incorporate social networks data and spillover information; these same authors provided feedback on the proposed analysis plan. JL and YL conducted analyses for the paper. All authors helped revise the drafts of the paper and approved the final manuscript.

## 9. Compliance with ethical standards

**Funding:** The authors received funding from the CDC via grant TZ/UG.08.0147, the President's Emergency Plan for AIDS Relief - Tanzania, and the UTHealth Innovation for Cancer Prevention Research Training Program Pre-doctoral Fellowship (Cancer Prevention and Research Institute of Texas grant RP160015).

**Conflict of interest:** The authors declare that they have no conflict of interest.

**Ethical approval:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. IRB approval was obtained from Harvard Medical School, the Muhimbili University of Health and Allied Sciences, and the National Institute for Medical Research (NIMR) in Tanzania. Particular attention was paid to eliminate the risk of inadvertent disclosure of CA HIV status to NMs.

**Informed consent:** Informed consent was obtained from all individual participants included in the study.

**Disclaimer:** The content is solely the responsibility of the authors and does not necessarily represent the official views of the Cancer Prevention and Research Institute of Texas.

1  
2  
3  
4       **Data sharing statement:** There are no additional data available. As the data contain enough information  
5 to potentially-uniquely identify specific participants, in the sensitive context of HIV, we have chosen not to  
6 make the data available.  
7  
8

## 9 10 **10. References**

- 11  
12 [1] Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and  
13 national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the  
14 Global Burden of Disease Study 2016. *The Lancet*. 2017;390(10100):1151–1210.  
15  
16 [2] Crepaz N, Tungol-Ashmon MV, Higa DH, Vosburgh W, Mullins MM, Barham T, et al. A systematic  
17 review of interventions for reducing HIV risk behaviors among people living with HIV in the  
18 United States, 1988-2012. *AIDS*. 2014 mar;28(5):633–56. Available from:  
19 <http://www.ncbi.nlm.nih.gov/pubmed/24983541>.  
20  
21 [3] Arnold EA, Kegeles SM, Pollack LM, Neilands TB, Cornwell SM, Stewart WR, et al. A Randomized  
22 Controlled Trial to Reduce HIV-Related Risk in African American Men Who Have Sex with Men and  
23 Women: the Bruthas Project. *Prevention Science*. 2018;p. 1–11.  
24  
25 [4] Garcia de Olalla P, Knobel H, Carmona A, Guelar A, Lopez-Colomes J, Cayla J. Impact of adherence  
26 and highly active retroviral therapy on survival in HIV-infected patients. *Journal of Acquired Immune  
27 Deiciency Syndromes*. 2002;30(1):105–10.  
28  
29 [5] Joint United Nations Programme on HIV/AIDS and Joint United Nations Programme on HIV/Aids and  
30 others. 90-90-90: an ambitious treatment target to help end the AIDS epidemic. Geneva: UNAIDS.  
31 2014;.  
32  
33 [6] Fisher JD, Fisher WA. Changing AIDS-risk behavior. *Psychological bulletin*. 1992;111(3):455.  
34  
35 [7] McNeill C, George N, Glover R. An Evaluation of Sisters Informing Healing Living Empowering:  
36 Increasing HIV Knowledge Among African American Adolescent Females Using an Evidence-Based  
37 HIV Prevention Intervention. *Journal of Doctoral Nursing Practice*. 2017;10(1):4–10.  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

- 1  
2  
3 [8] Carey MP, Braaten LS, Maisto SA, Gleason JR, Forsyth AD, Durant LE, et al. Using information,  
4 motivational enhancement, and skills training to reduce the risk of HIV infection for low-income urban  
5 women: A second randomized clinical trial. *Health Psychology*. 2000;19(1):3.  
6  
7  
8 [9] Sinclair B, McConnell M, Green DP. Detecting spillover effects: Design and analysis of multilevel  
9 experiments. *American Journal of Political Science*. 2012;56(4):1055–1069.  
10  
11  
12 [10] Benjamin-Chung J, Arnold BF, Berger D, Luby SP, Miguel E, Colford Jr JM, et al. Spillover effects in  
13 epidemiology: parameters, study designs and methodological considerations. *International journal of*  
14 *epidemiology*. 2017;47(1):332–347.  
15  
16  
17 [11] Truelove HB, Carrico AR, Weber EU, Raimi KT, Vandenberg MP. Positive and negative spillover of  
18 pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental*  
19 *Change*. 2014;29:127–138.  
20  
21  
22 [12] Ghosh D, Krishnan A, Gibson B, Brown SE, Latkin CA, Altice FL. Social network strategies to address  
23 HIV prevention and treatment continuum of care among at-risk and HIV-infected substance users: a  
24 systematic scoping review. *AIDS and Behavior*. 2017;21(4):1183–1207.  
25  
26  
27 [13] Benjamin-Chung J, Abedin J, Berger D, Clark A, Jimenez V, Konagaya E, Tran D, Arnold BF,  
28 Hubbard AE, Luby SP, and Miguel E, 2017. Spillover effects on health outcomes in low-and middle-  
29 income countries: a systematic review. *International journal of epidemiology*, 46(4), pp.1251-1276.  
30  
31  
32 [14] Self S, and Grabowski R, 2018. Factors influencing knowledge of HIV/AIDS in Nepal: role of  
33 socioeconomic interactions. *Journal of Social and Economic Development*, 20(1), pp.17-191.  
34  
35  
36 [15] Banerjee A, La Ferrara E, and Orozco-Olvera VH, 2019. The entertaining way to behavioral change:  
37 Fighting HIV with MTV. *The World Bank*.  
38  
39  
40 [16] Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *The New*  
41 *England journal of medicine*. 2007;357(4):370–379.  
42  
43  
44 [17] Perkins JM, Subramanian SV, Christakis Na. Social networks and health: A systematic review of  
45 sociocentric network studies in low- and middle-income countries. *Social science & medicine*. 2014  
46 aug;125(1):60–78. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25442969>.  
47  
48  
49  
50  
51  
52

- 1  
2  
3 [18] Coleman J, Katz E, Menzel H. The Diffusion of an Innovation among Physicians. *Sociometry*.  
4 1957;20(4):253–270.  
5  
6  
7 [19] VanderWeele TJ. Inference for influence over multiple degrees of separation on a social network.  
8 *Statistics in medicine*. 2013 feb;32(4):591–6; discussion 597–9. Available from:  
9 <http://www.ncbi.nlm.nih.gov/pubmed/23341081>.  
10  
11  
12 [20] Lienert J, Marcum CS, Finney J, Reed-Tsochas F, Koehly L. Social influence on 5-year survival in a  
13 longitudinal chemotherapy ward co-presence network. *Network Science*. 2017;5(3):308–327.  
14  
15 [21] Hughes JJ. Paying injection drug users to educate and recruit their peers: why participant-driven  
16 interventions are an ethical public health model. *Quality management in health care*. 1999;7(4):4–12.  
17  
18 [22] Ssali S, Wagner G, Tumwine C, Nannungi A, Green H. HIV Clients as Agents for Prevention: A Social  
19 Network Solution. *AIDS research and treatment*. 2012;2012:815823.  
20  
21 [23] Cislaghi B, Denny EK, Cisse´ M, Gueye P, Shrestha B, Shrestha PN, et al. Changing Social Norms: the  
22 Importance of “Organized Diffusion” for Scaling Up Community Health Promotion and Women  
23 Empowerment Interventions. *Prevention Science*. 2019;p. 1–11.  
24  
25 [24] Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden popula-  
26 tions for HIV surveillance. *Aids*. 2005;19:S67–S72.  
27  
28 [25] Smith R, Rossetto K, Peterson BL. A meta-analysis of disclosure of one’s HIV-positive status, stigma  
29 and social support. *AIDS care*. 2008;20(10):1266–1275.  
30  
31 [26] Watters JK, Biernacki P. Targeted sampling: options for the study of hidden populations. *Social*  
32 *problems*. 1989;36(4):416–430.  
33  
34 [27] Benjamin-chung J, Abedin J, Berger D, Clark A, Jimenez V, Konagaya E, et al. Spillover effects on  
35 health outcomes in low- and middle-income countries: a systematic review. *International Journal of*  
36 *Epidemiology*. 2017;(September):1–26.  
37  
38 [28] Neema S, Atuyambe LM, Otolok-Tanga B, Twijukye C, Kambugu A, Thayer L, et al. Using a clinic  
39 based creativity initiative to reduce HIV related stigma at the Infectious Diseases Institute, Mulago  
40 National Referral Hospital, Uganda. *African health sciences*. 2012;12(2):231–239.  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52



- 1  
2  
3  
4 [29] Smith Fawzi MC, Siril H, Li N, McAdam K, Ainebyona D, McAdam E, et al. Agents of Change:  
5 stepped wedge randomized controlled trial of the NAMWEZA intervention with people living with  
6 HIV and members of their social networks. *BMJ Global Health*. 2019;.
- 7  
8  
9 [30] Kaaya S, Arthur G, Siril H, Lienert J, Li N, McAdam K, et al. Leadership of PLWH as Agents of  
10 Change: comparing HIV-related risk behavior of people attending ART clinics in Dar es Salaam with  
11 members of their social networks. In: 20th International AIDS Conference; 2014..
- 12  
13  
14 [31] Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC medical research*  
15 *methodology*. 2006;6(1):54.
- 16  
17  
18 [32] Zou J, Yamanaka Y, John M, Watt M, Ostermann J, Thielman N. Religion and HIV in Tanzania:  
19 influence of religious beliefs on HIV stigma, disclosure, and treatment attitudes. *BMC public health*.  
20 2009;9(1):75.
- 21  
22  
23 [33] McCarthy B. Appreciative Inquiry: an alternative to behaviour management. *Dementia*.  
24 2017;16(2):249–253.
- 25  
26  
27 [34] Siril HN, Kaaya SF, Fawzi MKS, Mtisi E, Somba M, Kilewo J, et al. CLINICAL outcomes and loss to  
28 follow-up among people living with HIV participating in the NAMWEZA intervention in Dar es  
29 Salaam, Tanzania: a prospective cohort study. *AIDS research and therapy*. 2017;14(1):18.
- 30  
31  
32 [35] Carey MP, Schroder KE. Development and psychometric evaluation of the brief HIV Knowledge  
33 Questionnaire. *AIDS education and prevention*. 2002;14(2):172–182.
- 34  
35  
36 [36] Haile BJ, Chambers JW, Garrison JL. Correlates of HIV knowledge and testing: Results of a 2003  
37 South African survey. *Journal of Black Studies*. 2007;38(2):194–208.
- 38  
39  
40 [37] Aarø LE, Breivik K, Klepp KI, Kaaya S, Onya HE, Wubs A, et al. An HIV/AIDS knowledge scale for  
41 adolescents: item response theory analyses based on data from a study in South Africa and Tanzania.  
42 *Health education research*. 2011;26(2):212–224.
- 43  
44  
45 [38] Peltzer K, Matseke G, Mzolo T, Majaja M. Determinants of knowledge of HIV status in South Africa:  
46 results from a population-based HIV survey. *BMC public health*. 2009;9(1):174.
- 47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 [39] Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. *Pharmaceutical Statistics: The*  
4 *Journal of Applied Statistics in the Pharmaceutical Industry*. 2009;8(1):50-61.  
5  
6  
7 [40] Easterly W. How the millennium development goals are unfair to Africa. *World development*.  
8 2009;37(1):26–35.  
9  
10  
11 [41] Barbarin OA, Khomo N. Indicators of economic status and social capital in South African townships:  
12 What do they reveal about the material and social conditions in families of poor children? *Childhood*.  
13 1997;4(2):193–222.  
14  
15  
16 [42] Smith JA, McPherson M, Smith-Lovin L. Social distance in the United States: Sex, race, religion, age,  
17 and education homophily among confidants, 1985 to 2004. *American Sociological Review*. 2014  
18 Jun;79(3):432-56.  
19  
20  
21 [43] Lesko CR, Edwards JK, Cole SR, Moore RD, Lau B. When to censor? *American journal of epidemi-*  
22 *ology*. 2017;187(3):623–632.  
23  
24  
25 [44] Hernán MA. The hazards of hazard ratios. *Epidemiology (Cambridge, Mass.)*. 2010;21(1):13.  
26  
27  
28 [45] VanderWeele T. *Explanation in causal inference: methods for mediation and interaction*. Oxford  
29 University Press; 2015.  
30  
31  
32 [46] Hong, G., 2015. *Causality in a social world: Moderation, mediation and spill-over*. John Wiley & Sons.  
33  
34  
35 [47] Forastiere L, Mealli F, Vanderweele TJ. Identification and Estimation of Causal Mechanisms in  
36 Clustered Encouragement Designs: Disentangling Bed Nets Using Bayesian Principal Stratification  
37 Identification and Estimation of Causal Mechanisms in Clustered Encouragement Designs : Disentan-  
38 gling Be. *Journal of the American Statistical Association*. 2016;111(514):510–525. Available from:  
39 <http://dx.doi.org/10.1080/01621459.2015.1125788>.  
40  
41  
42  
43 [48] Imai K, Jiang Z. Identification and sensitivity analysis of contagion effects in randomized placebo-  
44 controlled trials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2019.  
45  
46  
47 [49] Shalizi CR, Thomas AC. Homophily and contagion are generically confounded in observational social  
48 network studies. *SMR*. 2012;40(2):211–39.  
49  
50  
51  
52

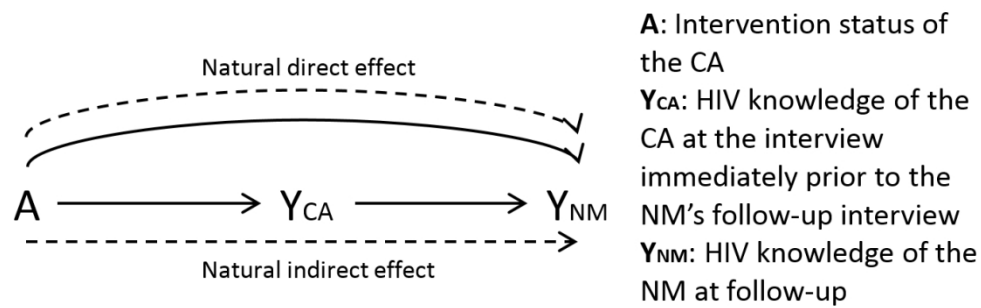
- 1  
2  
3 [50] De La Haye K, Robins G, Mohr P, Wilson C. Homophily and contagion as explanations for weight  
4 similarities among adolescent friends. *Journal of Adolescent Health*. 2011;49(4):421–427.  
5  
6  
7 [51] Vanderweele TJ, Hong G, Jones SM, Brown JL. Mediation and spillover effects in group-randomized  
8 trials: a case study of the 4Rs educational intervention. *Journal of the American Statistical Association*.  
9 2013;108(502):469–482. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23997375>.  
10  
11  
12 [52] Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal  
13 interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological*  
14 *methods*. 2013;18(2):137.  
15  
16  
17 [53] Leung MP. Treatment and spillover effects under network interference. *Review of Economics and*  
18 *Statistics*. 2016;p. 1–42.  
19  
20  
21 [54] Latkin CA, Davey-Rothwell MA, Knowlton AR, Alexander KA, Williams CT, Boodram B. Social  
22 network approaches to recruitment, HIV prevention, medical care, and medication adherence. *Journal*  
23 *of Acquired Immune Deiciency Syndromes*. 2013;63:1–9.  
24  
25  
26 [55] Galea S, Tracy M. Participation rates in epidemiologic studies. *Annals of epidemiology*.  
27 2007;17(9):643–653.  
28  
29  
30 [56] Berkman LF. Social support, social networks, social cohesion and health. *Social work in health care*.  
31 2000 Sep 26;31(2):3-14.  
32  
33  
34 [57] Langford CPH, Bowsher J, Maloney JP, Lillis PP. Social support: a conceptual analysis. *Journal of*  
35 *advanced nursing*. 1997;25(1):95–100.  
36  
37  
38 [58] Schneider JA, Zhou AN, Laumann EO. A new HIV prevention network approach: sociometric peer  
39 change agent selection. *Social science & medicine*. 2015;125:192–202.  
40  
41  
42 [59] He P, Eriksson F, Scheike TH, Zhang MJ. A proportional hazards regression model for the  
43 subdistribution with covariates-adjusted censoring weight for competing risks data. *Scandinavian*  
44 *Journal of Statistics*. 2016 Mar;43(1):103-22.  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## 11. Figure Captions

Fig 1. Schematic of natural direct effect (NDE) and natural indirect effect (NIE). The NDE indicates the increase in NMs' HIV knowledge happens without a concomitant increase in their CA's HIV knowledge. The NIE, on the other hand, indicates that the increase in NMs' HIV knowledge is mediated by their CA's HIV knowledge increasing. Solid lines indicate paths of causality between variables. Dashed lines represent the line or lines composing the effect of interest.

For peer review only

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



## STROBE (Strengthening The Reporting of OBServational Studies in Epidemiology) Checklist

A checklist of items that should be included in reports of observational studies. You must report the page number in your manuscript where you consider each of the items listed in this checklist. If you have not included this information, either revise your manuscript accordingly before submitting or note N/A.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).

Section and Item	Item No.	Recommendation	Reported on Page No.
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	
<b>Introduction</b>			
Background/Rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
<b>Methods</b>			
Study Design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed  <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	

Section and Item	Item No.	Recommendation	Reported on Page No.
Data Sources/ Measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study Size	10	Explain how the study size was arrived at	
Quantitative Variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	
Statistical Methods	12	(a) Describe all statistical methods, including those used to control for confounding	
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive Data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome Data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	

Section and Item	Item No.	Recommendation	Reported on Page No.
Main Results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other Analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key Results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
<b>Other Information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Once you have completed this checklist, please save a copy and upload it as part of your submission. DO NOT include this checklist as part of the main manuscript document. It must be uploaded as a separate file.**



# BMJ Open

## Evaluating spillover of HIV knowledge from study participants to their network members in a stepped-wedge behavioral intervention in Tanzania

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-033759.R5
Article Type:	Original research
Date Submitted by the Author:	06-Aug-2020
Complete List of Authors:	Lienert, Jeffrey; University of Pennsylvania Perelman School of Medicine, Center for Health Care Innovation; Philadelphia VAMC, CHERP Fawzi, Mary C; Harvard Medical School, Global Health and Social Medicine McAdam, Keith; London School of Hygiene and Tropical Medicine, Epidemiology Kaaya, Sylvia; Muhimbili University of Health and Allied Sciences, School of Medicine Liu, Yuanyuan Todd, Jim; London School of Hygiene and Tropical Medicine, Population Health Andrew, Irene; Management and Development for Health, Onnela, Jukka Pekka; Harvard T.H. Chan School of Public Health, Biostatistics
<b>Primary Subject Heading</b>:	HIV/AIDS
Secondary Subject Heading:	Global health, Public health
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, STATISTICS & RESEARCH METHODS

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3  
4  
5 Evaluating spillover of HIV knowledge from study participants to their  
6 network members in a stepped-wedge behavioral intervention in Tanzania  
7  
8

9 Jeffrey Lienert<sup>a,b,j</sup>, Mary C. Smith Fawzi<sup>c</sup>, Keith McAdam<sup>d</sup>, Sylvia Kaaya<sup>e</sup>, Yuanyuan Liu<sup>i</sup>, Jim Todd<sup>g</sup>,  
10 Irene Andrew Lema<sup>f</sup>, Jukka-Pekka Onnela<sup>h</sup>

11  
12 <sup>a</sup>Perelman School of Medicine, University of Pennsylvania <sup>b</sup>Center for Health  
13 Equity Research and Promotion, Philadelphia VA <sup>c</sup>Department of Global Health  
14 and Social Medicine, Harvard Medical School <sup>d</sup>Department of Clinical Tropical  
15 Medicine, London School of Hygiene and Tropical Medicine <sup>e</sup>Department of Psychiatry and  
16 Mental Health, Muhimbili University of Health and Allied Sciences <sup>f</sup>Management and  
17 Development for Health <sup>g</sup>Department of Population Health, London School of  
18 Hygiene and Tropical Medicine <sup>h</sup>Department of Biostatistics, Harvard T. H. Chan School of  
19 Public Health <sup>i</sup>University of Texas School of Public Health  
20 <sup>j</sup>Corresponding Author: Jeffrey Lienert, jeffrey.lienert@penntmedicine.upenn.edu  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

52 Preprint submitted to BMJ Open

3 August, 2020

## Abstract

**Objectives:** We aim to describe the social network members of participants of a behavioral intervention, and examine how the effects of the intervention may spillover among network members.

**Design:** Secondary analysis of a step-wedge randomized controlled trial.

**Setting:** Change Agents (CAs) were recruited from waiting rooms of HIV treatment facilities in Dar es Salaam, Tanzania, and their Network Members (NMs) were recruited directly by CAs.

**Participants:** We enrolled 662 CAs in an HIV behavioral intervention. They, along with 710 of their NMs, completed baseline and follow-up interviews from 2011 to 2013.

**Primary and Secondary Outcomes:** The primary outcome of this study was change in NMs' HIV knowledge, and the secondary outcome was whether the NM was lost to follow-up.

**Results:** At baseline, many characteristics were different between NMs and CAs. We found a number of NM characteristics significantly associated with follow-up of NMs, particularly female gender (OR=1.64, 95% CI: 1.02,2.63) and HIV knowledge (OR=20.0, 95% CI: 3.70,125); only one CA variable was significantly associated with NM follow-up: having a private source of water (OR=2.17, 95% CI: 1.33,3.57). The 14.2% increase in NMs' HIV knowledge was largely due to CAs feeling empowered to pass on prior knowledge, rather than transmitting new knowledge to their NMs.

**Conclusions:** Characteristics of social network members of persons living with HIV PLH may play a role in study retention. Additionally, the HIV knowledge of these NMs increased largely as a function of CA participation in the intervention, suggesting that intervening among highly-connected individuals may maximize benefits to the potential population for whom spillover can occur.

*Keywords:*

mediation, natural direct effect, natural indirect effect, follow-up, retention, hidden population

## 1. Strengths and Limitations of This Study

- The recruitment method, focused on social network members of persons living with HIV, allowed us to access an at-risk population typically not easily-accessible via other means, such as direct approach by the study team rather than someone known to the participant.
- The study design of the trial in which this study is nested was an ideal situation in which to employ a mediation analysis that informs our understanding of how this and similar intervention effects may spillover in a population.
- The greater-than-ideal dropout rate of NMs was both a strength in that it allowed us to examine factors associated with dropout, but also a limitation, in that the potential of differential dropout by unmeasured factors may have biased some of our results.

## 2. Introduction

Despite recent decreases in mortality, HIV/AIDS is currently the ninth leading cause of years of life lost (YLL) globally [1]. Research into eliminating HIV has focused on two fronts: biological and behavioral, which combined have decreased the years of life lost due to HIV by an estimated 51% [1]. Behavioral interventions seek to curb transmission of HIV through reduction of risk behavior including safer sexual practices and reducing injection drug use [2, 3]. Behavioral change also impacts the effectiveness of the biological methods, given that lower adherence is known to often compromise their effectiveness [4]. However, neither biological nor behavioral methods are fully effective without the knowledge of these treatments and preventative behaviors. It is therefore important to increase HIV-related knowledge as a necessary, but not sufficient, step towards achieving the 90-90-90 treatment goals (90% diagnosis, 90% antiretroviral therapy, and 90% viral suppression among the treated) and ultimately eliminate HIV [5].

The importance of adequate HIV knowledge has been recognized for a period of time, as it is often a prerequisite for behavioral change [6]. It should be noted that increased knowledge does not necessarily translate into behavior change, as other factors such as motivation and skills also play a role [6]. Despite this, researchers conducted trials among Persons Living with HIV (PLH) which led to increased HIV knowledge [7, 8]. These interventions were group discussions and teaching led by nurses, and motivational

1  
2  
3 interviewing, respectively. Researchers also found that the intervention reduced HIV-risk behaviors  
4 concomitantly with an increase in HIV knowledge [8].  
5

6  
7 1. Knowledge gained by participants in these trials can also be freely shared with members of their  
8 social network, in what is known as a spillover effect [9]. Specifically, a spillover effect (also known as an  
9 indirect or disseminated effect) is one person's exposure affecting another's outcome [10]. In the context of  
10 a social network spillover intervention, it corresponds to an individual who was unexposed to an intervention  
11 changing their behavior because they were socially-connected to an individual who did receive the  
12 intervention. This is distinct from what is sometimes called behavioral spillover, where changes in a person's  
13 behavior affect other behaviors of that same person [11]. For injection drug users, HIV prevention  
14 educational interventions were demonstrated to have spillover effects of HIV prevention education, and  
15 subsequent reduced rates of risky behaviors [12]. Studies have also used proxy variables for social network  
16 ties such as inviting social network members to watch educational programming [13] or time spent shopping  
17 at the market [14] to evaluate spillover effects for HIV knowledge, generally finding evidence for spillover.  
18 However, spillover in HIV knowledge between known social network ties generally remains understudied,  
19 particularly in sub-Saharan Africa [15]. We therefore aim to determine whether social network members  
20 those receiving an HIV behavioral/knowledge intervention also increase their HIV knowledge.  
21  
22  
23  
24  
25  
26  
27  
28  
29

30 New knowledge can come from a variety of sources, one of the most important of which is a person's  
31 social network [16, 17]. Social networks are of particular import because new knowledge can lead to  
32 cascades of behavior change, where people subsequently educate those in their social network, in what is  
33 known as social influence [18, 19, 20]. This has been examined in participant-driven interventions, where  
34 initially-recruited participants educate members of their social network one-on-one [21]. Characteristics  
35 such as knowing the HIV status of network members has been shown to be the most important predictor of  
36 engaging in prevention advocacy [22]. Work on diffusion through social networks, how a belief or behavior  
37 can be "contagious" within a network, has shown that spreading intervention effects beyond the initial study  
38 population can improve the cost-effectiveness of these interventions [23]. These findings imply that certain  
39 aspects of knowledge or behavior may spread more or less efficiently through networks comprising  
40 individuals with specific characteristics, which may need to be accounted for in network interventions. For  
41 instance, networks comprising many at-risk individuals who are HIV-negative may not be as receptive to  
42 change in behaviors as networks comprising a mix of those with and without HIV. Additionally, the  
43 networks of PLH are often difficult networks to ascertain [24], due to the continued stigma of HIV and AIDS  
44  
45  
46  
47  
48  
49  
50  
51  
52

1  
2  
3 in many settings [25]. Because of this, if a PLH or person at-risk of acquiring HIV does not want to  
4 participate in an investigator-initiated intervention, there is little recourse other than information transmitted  
5 via social networks, or targeted sampling techniques which are not always effective (e.g. Respondent Driven  
6 Sampling) [26]. This is particularly important in low- and middle-income countries, as a recent systematic  
7 review found only 54 studies researching spillover effects in these settings (out of approximately 750) [27].  
8 Therefore, understanding exactly how information spreads from participants in an intervention to members  
9 of their social network, who may be largely inaccessible via other means, is important for reaching the  
10 greatest number of people about HIV prevention. Understanding what makes these persons different from  
11 those who receive the intervention themselves is important, as it may point to ways in which to increase  
12 enrollment of these populations.  
13

14  
15 Based on the above gaps in the literature, we conduct a study on network members of PLH enrolled in a  
16 behavior change intervention [28]. The trial recruited PLH to serve as Change Agents (CAs) and to reach  
17 out to their social network members (NMs) about knowledge of HIV and safer sexual practices [29]. Our  
18 goals in this study were threefold: 1) to understand how the NMs differed from their CAs, 2) understand  
19 correlates of dropout for the NMs, and 3) to understand by what method HIV knowledge transferred to the  
20 NMs from the CAs. Understanding how the information and behaviors are shared within social networks  
21 will allow HIV researchers and others to take advantage of this knowledge and improve upon prevention  
22 interventions in the future.  
23  
24

### 25 3. Methods

#### 26 3.1. Study Population

27  
28 We analyze social network data from the Agents of Change trial [30], which was a stepped-wedge  
29 randomized controlled trial [31] that enrolled PLH to become Change Agents (CAs) by informing members  
30 of their social network (NMs) about knowledge of HIV and safer sexual practices. Here, we define CAs  
31 based on the *potential* for PLH to become so - by self-selecting into the study, PLH identify themselves as  
32 potential CAs. Although we refer to them as 'CAs' throughout, participants in the trial enrolled with varying  
33 levels of ability to act as a Change Agent. Through receiving the *NAMWEZA* intervention, we hypothesize  
34 that CAs will be able to truly self-actualize and subsequently act as Change Agents in their community.  
35

36  
37 CAs were recruited from the waiting rooms of HIV care and treatment centers in Dar es Salaam, Tanza-  
38 nia, and we received written consent from each CA. Participants completed a baseline questionnaire and  
39

1  
2  
3 were randomized to one of three waves in which to receive the intervention. At baseline, participants were  
4 also asked to recruit up to three members of their social networks who they felt were at particularly high risk  
5 of contracting or spreading HIV. We obtained written consent from these nominated network members (NMs).  
6 NMs could be either HIV positive or negative, and they were given a baseline survey. The NM was only  
7 aware the CA was a participant in the intervention if the CA shared this information with them, which  
8 many did not due to HIV-related stigma [32]. Each CA therefore formed a CA-NM dyad with each NM they  
9 recruited, and if they recruited more than one NM, formed a set of CA-NM dyads with a common CA.

10  
11 As fits a stepped-wedge RCT, all CAs eventually received the intervention, but were randomized to  
12 *when* they received it. These waves each lasted 12 weeks, at which point the next wave began and another  
13 group of CAs received the intervention. Within each wave, the intervention comprised 10 weekly structured  
14 sessions aimed at empowering PLH to become HIV prevention change agents in their communities. The  
15 sessions aimed to increase psychosocial and communication skills through an Appreciative Inquiry approach  
16 [33]. Within one month of each wave of the intervention, CAs were given follow-up surveys. Across all  
17 waves, the interventions lasted from November 2010 to January 2014, and the final interviews were  
18 conducted in March 2014. For more information on the study design, we direct interested readers to Smith-  
19 Fawzi et al., 2019 [29].

20  
21 The NMs did not receive any intervention at any point during the study. Rather, their intervention status  
22 flowed from the intervention status of their CAs. Due to operational difficulties, each NM was surveyed two  
23 times during the study: baseline and after the first wave, rather than baseline and one after each wave. In  
24 this way, all demographic and contextual variables were measured at baseline. We use this interview of NMs  
25 as the division between “exposed” and “unexposed”. At the time of an NM’s follow-up interview, their  
26 respective CA may or may not have undergone the intervention. In other words, the CAs randomized to  
27 receive *NAMWEZA* during the first wave would have potentially indirectly exposed their NMs to the  
28 intervention when the NM completed their follow-up questionnaire after Wave 1. Therefore, the NMs were  
29 divided into “exposed” (N=381) and “unexposed” (N=329) groups based on whether their respective CA  
30 was randomized into receiving *NAMWEZA* during the first wave or not. The CAs always completed their  
31 Wave 1 follow-up interview before their NMs were invited to complete their Wave 1 follow-up interview.

32  
33 Inclusion criteria for the follow-up analyses were for CAs who completed at least one follow-up inter-  
34 view and had at least one NM who completed at least one follow-up questionnaire. In this way, outcomes  
35 could be computed for the CA and NM of each dyad. NMs approached by a CA, but who did not participate



1  
2  
3 in the study were not recorded since it was not feasible to obtain this information from study participants  
4 themselves. As we lost some CAs and NMs to follow up, we completed our analyses without their data,  
5 assuming it to be Missing Completely at Random (MCAR). During this study, there was little loss-to-follow-  
6 up among the CAs (< 10%), but much higher loss among the NMs (36.8%) [34]. Given an NM or CA was  
7 not lost to follow-up, complete information was available on all additional variables, including exposure,  
8 outcome, and covariates. In sum, our sample comprises 662 CAs and 710 NMs, meaning each CA recruited  
9 1.07 NMs on average out of a possible 3, and 44 CAs nominated at least two NMs.  
10  
11  
12  
13  
14  
15

### 16 *13.2. HIV knowledge*

17  
18 To assess HIV knowledge of CAs and NMs, we used the brief HIV knowledge questionnaire [35]. This  
19 scale comprises 18 questions, which focus on an individual's knowledge of how HIV can spread and other  
20 characteristics of the virus and AIDS. This knowledge is crucial for an individual's subsequent safe-sex  
21 practices to reduce the risk of transmission. The original population comprised three different groups: two  
22 groups of low-income women, and one of women and men receiving psychiatric treatment. In these popula-  
23 tions, questions on the measure had a Chronbach's alpha of 0.78. This instrument has been used previously  
24 in sub-Saharan Africa and has demonstrated reasonable reliability of 0.75 Cronbach's alpha in South Africa  
25 among a convenience sample of 429 members of the African Methodist Episcopal church [36]. It has also  
26 been translated to Swahili, with only minor differential item functioning [37]. This indicates that the  
27 measure performs adequately in other, similar populations. In the present study the Cronbach's alpha is  
28 0.71. Because the average baseline score of HIV knowledge was 0.80, with a preponderance of scores of  
29 1.00, the main indicator of knowledge was not normally distributed, and therefore a continuous predictor  
30 was not ideal [38,39]. We therefore summarize this measure as "Complete HIV Knowledge", a dichotomous  
31 variable for whether the participant correctly answered all questions.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

### 42 *13.3. Demographic and contextual variables*

43  
44 In addition to HIV knowledge, we are interested in a set of demographic and contextual variables that  
45 may help explain some of the trends observed. In terms of demographic variables, we include age, sex,  
46 employment status, marital status, and self-identified HIV status. We also include education, which we di-  
47 chotomize at 7 years or more - a level commensurate with elementary education. We selected this cutoff  
48  
49  
50  
51  
52

1  
2  
3 because it coincides with the millennium development goal (MDG) of increasing primary education com-  
4 pletion [40].  
5  
6

7 Contextually, we included two additional variables: having a private source of water, and the number  
8 of persons sleeping in the participant's home. The first was also based on a MDG, and indicates participants  
9 with access to safe drinking water [40]. This is a proxy for the economic security of the participant. The  
10 number of persons sleeping in the participant's home is also a non-monetary indicator of their material and  
11 social resources [41]. These variables combined give a more thorough picture of the participant's economic  
12 status than employment alone.  
13  
14  
15

### 16 17 *13.4. Statistical Analysis* 18

19 The first analysis was to assess homophily, defined as the extent to which CAs and NMs were similar to  
20 one another in terms of a number of sociodemographic characteristics and HIV-related risk factors. Because  
21 CAs and NMs self-selected into their respective group, and only CAs were randomized (with their NMs  
22 being randomized along with them), differences between the groups were to be expected. However, we only  
23 examined homophily of baseline characteristics rather than of outcome, because comparison of outcomes  
24 between CAs and NMs would remove the benefits of randomization. Because CAs and NMs self-selected  
25 into the study and were not randomized to CA/NM status, we do not *a priori* expect them to be completely  
26 similar. In addition, multiple NMs could share a CA and would therefore not be independent due to the  
27 shared variation and latent characteristics of having the same CA. We therefore assessed statistical  
28 significance of homophily on the set of CA-NM dyads using a permutation test, a non-parametric test which  
29 has no distributional assumptions. For continuous variables, the difference between the CA and NM was  
30 calculated, and for categorical variables, whether the CA and NM were concordant or discordant was  
31 recorded. We then randomly permuted CA-NM ties (keeping number of ties per CA constant), and then  
32 recalculated the difference or percent concordant, respectively, 1,000 times. We then examine the percentile  
33 of the observed difference relative to the permuted differences [42]. Analyses were run using R v3.1.1.  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

44 To accomplish our second aim of understanding what was associated with NMs completing their follow-  
45 up interview, we fit a log-binomial regression to determine predictors of follow-up. In this regression, we  
46 use all the variables listed above, as well as whether the CA remained in the study for its full duration. In  
47 order to examine the association between these same variables and the time of follow-up, we also employed  
48 a Cox regression. In the Cox regression, start of follow-up was defined as the time at which the NM  
49  
50  
51  
52

1  
2  
3 completed their baseline interview. The outcome here was whether the NM completed a follow-up interview.  
4  
5 NMs who were lost to follow-up were censored at the time of their latest interview [43,44].  
6

7 Finally, since the trial showed beneficial effects on the HIV knowledge of the NMs following the inter-  
8 vention [30], we aim to elucidate exactly what caused the HIV knowledge of NMs to increase - either CAs  
9 gaining knowledge through the intervention and sharing it, or the CAs being empowered by the intervention  
10 to share existing knowledge. As the wedge in which the CA received the *NAMWEZA* intervention was  
11 randomized, we treat each NM as being randomized to exposure to *NAMWEZA* at the same time as their  
12 CA. This randomization scheme allows us to explore the spillover effect of CAs' HIV knowledge onto their  
13 respective NMs via a mediation analysis. These pathways represent different types of spillover effects: the  
14 exposure or outcome of one person affecting the outcome of another person.  
15  
16  
17  
18  
19

20 As shown by VanderWeele et al (2015), social network spillover effects in the case of dyadic  
21 relationships can be broken down into concepts from mediation analysis: direct and indirect effects (Figure  
22 1) [45]. This method has since been used for novel evaluations of spillover effects [46,47,48]. Although  
23 previous studies showed that this same type of analysis cannot be done on full network data, the data in this  
24 study consisted of only dyads, the CAs and their NMs, so in this case the method is appropriate [49,50].  
25  
26  
27  
28

29 The method parses social influence into direct and indirect effects. The natural indirect effect (NIE) is  
30 the effect by which the intervention changes the CA's HIV knowledge, which then changes the NM's HIV  
31 knowledge. The NIE is similar to the effect observed in many participant-driven interventions: an initial  
32 participant receives the intervention, increasing their knowledge, and they subsequently pass their increased  
33 knowledge to members of their social network [21]. The natural direct effect (NDE) is the effect of receiving  
34 the intervention has on an NMs outcome, irrespective of the CA's outcome (in this case HIV knowledge).  
35 For instance, this could occur if CAs begin with good knowledge of HIV, and the intervention empowers  
36 them to convey knowledge they already had to their NMs. Although the intervention does not increase their  
37 HIV knowledge, it is still useful to the CAs, as it empowers them to act as CAs in their community. In order  
38 to estimate these effects, the published SAS macro developed by VanderWeele and colleagues was used,  
39 and analyses were run using SAS v9.2 (SAS Institute, Cary, NC) [51,52]. In the models estimating the  
40 effect of the exposure on the mediator and estimating the effect of the mediator on the outcome,  
41 we adjusted for all the variables included in our logistic regression above. Although the  
42 randomization of the exposure minimized some potential bias, the loss-to-follow-up among the  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

1  
2  
3 CAs indicates that selection bias could remain a concern, so we control for the variables which  
4 may also impact loss-to-follow-up.  
5

6  
7 Importantly, this analysis requires a number of assumptions and applies to dyads only when these  
8 assumptions are met. One assumption of this analysis is that the dyads are independent, which is violated  
9 here; if a CA recruited more than one NM, the multiple CA-NM dyads involving the same CA would not  
10 be independent. To address this, we performed the analysis after randomly removing NMs until each CA  
11 had only a single NM. This resulted in removing 48 NMs, just 6.7% of the population. We found that the  
12 point estimates were nearly identical, but that the confidence intervals were slightly larger due to the reduced  
13 sample size. No coefficients changed from significant to non-significant in this analysis (data not shown).  
14  
15

16  
17 A second, related assumption is that of *partial interference*, that the effects in one cluster does not affect  
18 another cluster - here, one CA-NM dyad affecting another [53]. This could occur if two NMs of different  
19 CAs happen to know one-another outside of the study, one has a CA who was randomized to an earlier  
20 wedge, and shares what they know of it with the otherwise-unexposed NM. However, due to the large size  
21 of Dar es Salaam, and the number of HIV treatment clinics in which recruitment occurred, we expect few  
22 CAs or NMs to know one another outside of the study (other than NMs knowing the CA who recruited  
23 them), limiting the potential for partial interference.  
24  
25

26  
27 A third assumption of this analysis is that the outcomes of the CA and NM are independent conditional  
28 on the CA's exposure, or conditional on the CA's exposure and any confounding variables [45,46]. Because  
29 CA-NM pairs self-select and are not randomized, we do not expect these outcomes to be independent  
30 conditional of the CA's exposure, and so we adjust for additional variables to meet this assumption. In our  
31 analysis, we therefore adjust for all the variables used in the log-binomial regression.  
32  
33

34  
35 The data are not publicly available due to the sensitive nature of HIV infection status and the socially-  
36 networked nature of the data. Because the data include specific information on social ties, some of whom  
37 have not disclosed HIV sero-status to one another, the risks of individual identification and compromising  
38 HIV sero-status are greatly increased.  
39  
40  
41  
42  
43  
44  
45

#### 46 *13.5. Patient and Public Involvement*

47  
48 Patients were not involved in the design of this study. As part of the recruitment process, patients were  
49 instructed to recruit members of their social networks who they felt were at high risk of HIV infection - in  
50  
51  
52

this way participants could guide the recruitment of NMs to those most at-risk. We currently have no plans to disseminate the results of this study to participants.

#### 4. Results

The *NAMWEZA* study recruited 662 Change agents (CA), of whom 453 (68.4%) completed at least one follow-up questionnaire. These 662 CAs in turn recruited 710 network members (NM) who took the baseline questionnaire. Of the 710 NMs, 449 (63.2%) also completed a follow-up questionnaire (Table 1). At baseline, CAs were on average older than their NMs. More of the NMs were employed than their CAs (69.3% vs. 55.0%,  $p<0.001$ ), but were less likely to have at least 7 years of education (52.0% vs. 52.3%,  $p<0.001$ ). Only 12.3% of NMs were HIV-positive, compared to *all* CAs ( $p<0.001$ ). Complete data was obtained at baseline for all CAs and NMs.

Characteristic	Number of NMs (%) or Mean (SD) (N=710)	Number of CAs (%) or Mean (SD) (N=662)	P-value
Age	33.0 (11.1)	38.9 (9.7)	<0.001
Female	380 (53.7%)	349 (53.9%)	0.89
Employed	490 (69.3%)	356 (55.0%)	<0.001
At least 7 years education	369 (52.0%)	584 (82.3%)	<0.001
Complete HIV knowledge	638 (89.9%)	598 (90.4%)	0.65
Persons sleeping in home	5.00 (5.77)	3.87 (3.72)	<0.001
Married	373 (52.7%)	338 (51.1%)	0.56
HIV Positive	87 (12.3%)	662 (100%)	N/A
Private source of water	309 (43.7%)	263 (39.7%)	0.19

Table 1: Demographic characteristics at baseline, with the results of a permutation test for homophily. Specifically, for each CA- NM dyad, either a difference (for continuous variables) or concordance (for dichotomous variables) is calculated. For example, if a CA was 39 years old, and their NM was 25 years old, the difference would be 14 years old. If a CA was male and their NM was Female, the pair would be discordant for sex. CA-NM pairs were then randomly reshuffled, the edge-wise characteristics recalculated, and the observed difference compared to the distribution of randomized differences.

Risk ratios (RRs) obtained via log-binomial regression indicated that characteristics of both the CAs and the NMs significantly predicted loss to follow-up (Table 2). The NM being female (RR=1.44, 95% CI: 1.05,1.97), having complete HIV knowledge (RR=10, 95% CI: 2.33,42), being employed (RR=1.43, 95% CI: 1.08,1.89), and being married (RR=1.55, 95% CI: 1.03,2.33) were all significant predictors of increased odds of completing a follow-up interview. Each additional person sleeping in the home of the NM per room used for sleeping reduced the odds of follow-up (RR=0.85, 95% CI: 0.74,0.98) as did the NM living with

HIV (RR=0.40, 95% CI: 0.17,0.96). CA having a private water source was significantly associated with increased odds of the NM being followed-up (RR=2.07, 95% CI: 1.25,3.42), even after controlling for the NM having a private source of water. This was the only CA-specific variable that significantly predicted an NM's follow-up.

The Cox proportional hazard model showed similar results to the logistic regression, but with fewer significant results. Only the NM being married (HR=1.28, 95% CI: 1.01,1.64) and the NM living with HIV (HR=0.71, 95% CI: 0.51,0.99) significantly predicted the time-to-follow-up of the NMs. Additionally, all of the hazard ratios were closer to the null relative to the corresponding odds ratios.

We assess the mechanism of spillover effects on HIV knowledge, the main effect of the intervention on the NMs. The analysis showed that NMs who were exposed to intervention via their CA had an increase in HIV knowledge of 12.9% on average (95% CI: 0.06, 0.20) from a baseline percentage of 78% [26]. This was broken down into a natural indirect effect (NIE) of 0.6% (95% CI: -0.06%,2.0%), which is the effect the intervention had via the CA's HIV knowledge, and a natural direct effect (NDE) of 12.3% (95% CI: 6.1%,19.3%), which is the effect the CA's participation had on the NM's HIV knowledge, irrespective of the CA's HIV knowledge. In other words, of the 12.9% increase in NM's having complete HIV knowledge, 12.3% occurred without a concomitant increase in their CA's HIV knowledge. In other words, their HIV knowledge increase was not mediated by the increase in HIV knowledge of their CA. This did not change when we used only one NM per CA.

## 5. Discussion

Based on results from the NAMWEZA trial in Tanzania, we have shown a number of novel findings regarding the network members of HIV intervention participants, correlates of their retention in the study, and the evidence for spillover of HIV knowledge from CAs to NMs. These findings can inform the design of interventions in the future to maximally enroll and retain participants and ensure that intervention information is transmitted from the study participants to members of their social networks.

Characteristic (N=459)	Adjusted RR (95% CI)	Adjusted HR (95% CI)
Characteristics of NMs		
Gender (Female)	1.44* (1.05,1.97)	1.18 (0.94,1.50)
Difference in age of NM and CA (per year)	1.01 (0.99,1.03)	1.01 (0.995,1.02)
Age of NM	0.98 (0.95,1.02)	0.99 (0.97,1.04)
Complete HIV knowledge (vs. none)	10* (2.33,42)	2.20 (0.97,5.01)
Employed	1.43* (1.08,1.89)	1.15 (0.89,1.50)
Each additional person sleeping in home per room used for sleeping	0.85* (0.74,0.98)	0.92 (0.83,1.01)

Married	1.55* (1.03,2.33)	1.28* (1.01,1.64)
Living with HIV	0.40* (0.17,0.96)	0.71* (0.51,0.99)
Having a private source of water	0.89 (0.59,1.34)	0.97 (0.77,1.22)
Characteristics of CAs		
Gender (Female)	1.27 (0.76,2.08)	1.07 (0.84,1.37)
Complete HIV knowledge (vs. none)	0.36 (0.07,2.04)	0.64 (0.29,1.43)
Having a private source of water	2.07* (1.25,3.42)	0.97 (0.77,1.22)
Being employed	1.54 (0.95,2.50)	1.14 (0.90,1.43)
Being married	1.25 (0.77,2.04)	1.11 (0.88,1.43)
CA Lost to follow-up	1.06 (0.84,1.33)	1.02 (0.72,1.43)

Table 2: Results of multivariate log-binomial regression and Cox-proportional hazard models on the dichotomous outcomes of whether the NM completed a follow-up questionnaire, and the continuous outcome of time-to-completion of follow-up questionnaire, respectively. “\*” indicates significant at the  $p < 0.05$  level.

We found several significant differences between CAs and NMs they recruited. On average, CAs were older, less likely to be employed, and more educated. All of this suggests that by having CAs recruit from their social network, we were able to recruit a set of social network members different in multiple ways from those found in the waiting rooms in HIV treatment clinics (unsurprising, given that many NMs were not HIV-positive) [9]. This may in part be due to the fact that when recruiting NMs, CAs were instructed to recruit those who may be at particular risk of contracting or spreading HIV. Therefore, CAs likely did not pick random members of their social network, but those who were at high risk. This finding corroborates the claim by Latkin et al. (2013) that sampling via enumeration of network members by initial participants recruits a more diverse sample, as our sample of NMs was not composed of only those who were HIV-positive, but also many who were HIV-negative [54]. Additionally, the lower likelihood of employment of the CAs indicates they may have had more time for participating in the intervention [55]. This means that future studies may need to tailor their interventions to work with the schedules of employed persons to increase participation.

Although the study design potentially accessed a separate slice of the population than other methods vis-a'-vis the population of those at risk for contracting HIV, it may have been at least partially responsible for the large loss of follow-up of the NMs. It is important to note that more NM characteristics than CA characteristics were significantly predictive of follow-up in both the logistic and Cox proportional hazards models. The one CA characteristic which did predict NMs completing a follow-up interview, having a private source of water, was a proxy for CA's socioeconomic status. This may have been because CAs with greater resources may have had more time available to pass information to their NMs, retaining the NM's interest longer [56]. We also note that less HIV knowledge and living with HIV predicted reduced likelihood

1  
2  
3 of completing a follow-up questionnaire among NMs, which means that those who might have benefited  
4 most from spillover of the intervention were more likely to discontinue their involvement. This does not  
5 mean that they did not receive any spillover, only that it was not recorded. Our estimate of the magnitude of  
6 the spillover may therefore be biased towards the null. This is also problematic more generally for  
7 interventions of this nature as the very people the intervention aims to benefit may not stay with the program.  
8  
9

10  
11 The HIV knowledge gain experienced by the NMs was largely due to the NDE; i.e. knowledge spilled-  
12 over as a result of the CAs participating in NAMWEZA, not because the CAs' HIV knowledge increased,  
13 and they passed this new knowledge to their NM(s). This is most likely because CAs had a high average  
14 HIV knowledge score at baseline (80%), so there was less room for improvements in knowledge. Following  
15 the intervention, CAs' likelihood of complete HIV knowledge did not significantly increase. Therefore, what  
16 prompted the increased knowledge of the NMs was the CA becoming empowered through the intervention  
17 to pass on their existing knowledge to their NM [57].  
18  
19

20  
21 This finding is important for future interventions: spillover effects of this intervention will likely carry  
22 over only to those directly-connected to the CAs (opposed to spreading indefinitely in a snowball effect), as  
23 the most important thing in transmitting HIV knowledge is receiving the intervention, which only CAs  
24 receive. Because NMs do not actually receive the intervention, it is unlikely that they would feel empowered  
25 to become CAs themselves (particularly because they will not receive the *NAMWEZA* sessions), thus limiting  
26 the spread of the intervention. For CAs to increase the HIV knowledge of their NMs, they would only need  
27 to become empowered to share their information. However, for the NMs to increase the HIV knowledge of  
28 their own NMs, they would need to both increase their HIV knowledge, and become empowered to share it.  
29 This is less likely to happen than just becoming empowered, and so it is unlikely, though not impossible, that  
30 this effect would continue to spread in the population. This may give insight into how to design interventions  
31 in the future; if one wants to maximize the number of people who benefit, choosing CAs who form many  
32 bridging ties in the community would maximize the potential number of links by which spillover can occur  
33 [58]. Alternatively, interventions can be designed to be self-propagating; if CAs are empowered to deliver  
34 the intervention to others, changing their own NMs into future CAs, those new NMs-turned-CAs could then  
35 deliver the intervention to a second set of NMs, again empowering them to become CAs.  
36  
37

38  
39 Our study is not without limitations. First, there was considerable loss to follow-up of the NMs. Al-  
40 though this actually informed our analysis of the correlates of loss to follow-up, it meant that our analysis  
41 of spillover effects may be biased. Even though the exposure was randomized, the loss to follow-up can  
42  
43  
44  
45  
46  
47



1  
2  
3 result in selection bias if the NMs who left the study were systematically different from those who remained.  
4  
5 As we show, the NMs who dropped out were those who would have benefited the most from the intervention  
6  
7 because they were more vulnerable and at higher risk than those who did not drop out. Because they are  
8  
9 likely to have greater benefits from any spillover effects, we would expect this non-random loss to follow-up  
10  
11 to result in an underestimate of the impact of the intervention in the network members. Adjusting for  
12  
13 censoring weights may ameliorate this issue [59]. Second, our use of Hazard Ratios (HRs) has important  
14  
15 limitations: they are subject to selection bias, are sensitive to study period, and only provide one estimate  
16  
17 during the study [44]. Any of these limitations could affect this analysis, hence our use of logistic regression  
18  
19 as a primary analysis. However, they remain useful as a sensitivity analysis. Third, our data did not perfectly  
20  
21 fit the requirements of the causal mediation analysis: they were not entirely independent, since multiple NM-  
22  
23 CA dyads shared a CA. However, when we randomly removed dyads until there were no repeated CAs, the  
24  
25 results were qualitatively very similar, indicating that lack of independence did not unduly affect our results.  
26  
27 Fourth, although we were able to tease apart the direct and indirect effects, we are unable to determine the  
28  
29 mechanism of the natural direct effect; the data do not allow us to specify whether NMs increased knowledge  
30  
31 through speaking to their knowledgeable CA, through researching HIV on their own, or some other  
32  
33 mechanism. Future work will have to be done to examine these different pathways.

## 31 **6. Conclusions**

32  
33  
34 These results have implications for the potential scale-up of the *NAMWEZA* intervention, as well as  
35  
36 future studies and interventions that focus on behavioral interventions in social networks. First, our findings  
37  
38 of minimal similarity between CAs and their NMs indicate that this recruitment method allows us to enroll  
39  
40 participants from portions of the population that are not represented by the CAs alone. Coupling this with a  
41  
42 deeper understanding of the mental heuristics CAs used to select NMs (e.g. did CAs mentally search their  
43  
44 close or peripheral network for those at-risk of HIV), may lead to different strategies for recruitment and  
45  
46 retention, leading to stronger effects of behavioral interventions. The mediation analysis presents a  
47  
48 compelling picture of how best to ensure the benefits of interventions reach as many people beyond the  
49  
50 study participants as possible. Participation of CAs in the intervention resulted in positive effects on their  
51  
52 immediate network members' HIV knowledge regardless of how the CAs responded to the intervention.  
53  
54 While improvement in HIV knowledge may not necessarily translate to increased safe sex behavior, it can  
55  
56 be seen as a rate-limiting step towards reducing HIV transmission since adequate knowledge is usually

necessary for reducing risk behavior. Future work should examine the exact mechanisms of spillover discussed here, as that is an important clarification that could benefit future studies. Specifically, a similar setup to the study here combined with semi-structured interviews with CAs and NMs about their interactions with one another would help elucidate exactly how NMs increased their HIV knowledge (this population would likely need to comprise CAs who have disclosed their HIV status in order to prevent accidental disclosure). Future work should also examine whether increases in HIV knowledge translate to changes in behaviors which may increase one's risk of contracting HIV, particularly in Sub-Saharan Africa. The results presented herein may inform approaches for increasing participation and potentially conferring greater benefits related to spillover effects in future HIV behavioral interventions.

## 7. Acknowledgements

The authors would like to thank the study participants and staff for their time, hard work, and dedication. We would also like to thank Guy Harling for their comments on the manuscript.

## 8. Contributorship

JL, MCSF, and JPO designed the analysis plan and drafted the initial paper. JPL, MCSF, KM, SK, IAL and JT conducted the underlying randomized trial, and designed it to incorporate social networks data and spillover information; these same authors provided feedback on the proposed analysis plan. JL and YL conducted analyses for the paper. All authors helped revise the drafts of the paper and approved the final manuscript.

## 9. Compliance with ethical standards

**Funding:** The authors received funding from the CDC via grant TZ/UG.08.0147, the President's Emergency Plan for AIDS Relief - Tanzania, and the UTHealth Innovation for Cancer Prevention Research Training Program Pre-doctoral Fellowship (Cancer Prevention and Research Institute of Texas grant RP160015).

**Conflict of interest:** The authors declare that they have no conflict of interest.

**Ethical approval:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. IRB approval was obtained from Harvard Medical School, the Muhimbili University of Health and Allied Sciences, and the National Institute

for Medical Research (NIMR) in Tanzania. Particular attention was paid to eliminate the risk of inadvertent disclosure of CA HIV status to NMs.

**Informed consent:** Informed consent was obtained from all individual participants included in the study.

**Disclaimer:** The content is solely the responsibility of the authors and does not necessarily represent the official views of the Cancer Prevention and Research Institute of Texas.

**Data sharing statement:** There are no additional data available. As the data contain enough information to potentially-uniquely identify specific participants, in the sensitive context of HIV, we have chosen not to make the data available.

## 10. References

- [1] Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*. 2017;390(10100):1151–1210.
- [2] Crepaz N, Tungol-Ashmon MV, Higa DH, Vosburgh W, Mullins MM, Barham T, et al. A systematic review of interventions for reducing HIV risk behaviors among people living with HIV in the United States, 1988-2012. *AIDS*. 2014 mar;28(5):633–56. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24983541>.
- [3] Arnold EA, Kegeles SM, Pollack LM, Neilands TB, Cornwell SM, Stewart WR, et al. A Randomized Controlled Trial to Reduce HIV-Related Risk in African American Men Who Have Sex with Men and Women: the Bruthas Project. *Prevention Science*. 2018;p. 1–11.
- [4] Garcia de Olalla P, Knobel H, Carmona A, Guelar A, Lopez-Colomes J, Cayla J. Impact of adherence and highly active retroviral therapy on survival in HIV-infected patients. *Journal of Acquired Immune Deiciency Syndromes*. 2002;30(1):105–10.
- [5] Joint United Nations Programme on HIV/AIDS and Joint United Nations Programme on HIV/Aids and others. 90-90-90: an ambitious treatment target to help end the AIDS epidemic. Geneva: UNAIDS. 2014;.
- [6] Fisher JD, Fisher WA. Changing AIDS-risk behavior. *Psychological bulletin*. 1992;111(3):455.

- 1  
2  
3 [7] McNeill C, George N, Glover R. An Evaluation of Sisters Informing Healing Living Empowering:  
4 Increasing HIV Knowledge Among African American Adolescent Females Using an Evidence-Based  
5 HIV Prevention Intervention. *Journal of Doctoral Nursing Practice*. 2017;10(1):4–10.  
6  
7  
8  
9 [8] Carey MP, Braaten LS, Maisto SA, Gleason JR, Forsyth AD, Durant LE, et al. Using information,  
10 motivational enhancement, and skills training to reduce the risk of HIV infection for low-income urban  
11 women: A second randomized clinical trial. *Health Psychology*. 2000;19(1):3.  
12  
13  
14 [9] Sinclair B, McConnell M, Green DP. Detecting spillover effects: Design and analysis of multilevel  
15 experiments. *American Journal of Political Science*. 2012;56(4):1055–1069.  
16  
17  
18 [10] Benjamin-Chung J, Arnold BF, Berger D, Luby SP, Miguel E, Colford Jr JM, et al. Spillover effects in  
19 epidemiology: parameters, study designs and methodological considerations. *International journal of  
20 epidemiology*. 2017;47(1):332–347.  
21  
22  
23 [11] Truelove HB, Carrico AR, Weber EU, Raimi KT, Vandenberg MP. Positive and negative spillover of  
24 pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental  
25 Change*. 2014;29:127–138.  
26  
27  
28 [12] Ghosh D, Krishnan A, Gibson B, Brown SE, Latkin CA, Altice FL. Social network strategies to address  
29 HIV prevention and treatment continuum of care among at-risk and HIV-infected substance users: a  
30 systematic scoping review. *AIDS and Behavior*. 2017;21(4):1183–1207.  
31  
32  
33 [13] Benjamin-Chung J, Abedin J, Berger D, Clark A, Jimenez V, Konagaya E, Tran D, Arnold BF,  
34 Hubbard AE, Luby SP, and Miguel E, 2017. Spillover effects on health outcomes in low-and middle-  
35 income countries: a systematic review. *International journal of epidemiology*, 46(4), pp.1251-1276.  
36  
37  
38 [14] Self S, and Grabowski R, 2018. Factors influencing knowledge of HIV/AIDS in Nepal: role of  
39 socioeconomic interactions. *Journal of Social and Economic Development*, 20(1), pp.17-191.  
40  
41  
42 [15] Banerjee A, La Ferrara E, and Orozco-Olvera VH, 2019. The entertaining way to behavioral change:  
43 Fighting HIV with MTV. *The World Bank*.  
44  
45  
46 [16] Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *The New  
47 England journal of medicine*. 2007;357(4):370–379.  
48  
49  
50  
51  
52

- 1  
2  
3  
4 [17] Perkins JM, Subramanian SV, Christakis Na. Social networks and health: A systematic review of  
5 sociocentric network studies in low- and middle-income countries. *Social science & medicine*. 2014  
6 aug;125(1):60–78. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25442969>.  
7  
8  
9 [18] Coleman J, Katz E, Menzel H. The Diffusion of an Innovation among Physicians. *Sociometry*.  
10 1957;20(4):253–270.  
11  
12  
13 [19] VanderWeele TJ. Inference for influence over multiple degrees of separation on a social network.  
14 *Statistics in medicine*. 2013 feb;32(4):591–6; discussion 597–9. Available from:  
15 <http://www.ncbi.nlm.nih.gov/pubmed/23341081>.  
16  
17  
18 [20] Lienert J, Marcum CS, Finney J, Reed-Tsochas F, Koehly L. Social influence on 5-year survival in a  
19 longitudinal chemotherapy ward co-presence network. *Network Science*. 2017;5(3):308–327.  
20  
21  
22 [21] Hughes JJ. Paying injection drug users to educate and recruit their peers: why participant-driven  
23 interventions are an ethical public health model. *Quality management in health care*. 1999;7(4):4–12.  
24  
25  
26 [22] Ssali S, Wagner G, Tumwine C, Nannungi A, Green H. HIV Clients as Agents for Prevention: A Social  
27 Network Solution. *AIDS research and treatment*. 2012;2012:815823.  
28  
29  
30 [23] Cislighi B, Denny EK, Cisse' M, Gueye P, Shrestha B, Shrestha PN, et al. Changing Social Norms: the  
31 Importance of “Organized Diffusion” for Scaling Up Community Health Promotion and Women  
32 Empowerment Interventions. *Prevention Science*. 2019;p. 1–11.  
33  
34  
35 [24] Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden popula-  
36 tions for HIV surveillance. *Aids*. 2005;19:S67–S72.  
37  
38  
39 [25] Smith R, Rossetto K, Peterson BL. A meta-analysis of disclosure of one’s HIV-positive status, stigma  
40 and social support. *AIDS care*. 2008;20(10):1266–1275.  
41  
42  
43 [26] Watters JK, Biernacki P. Targeted sampling: options for the study of hidden populations. *Social*  
44 *problems*. 1989;36(4):416–430.  
45  
46  
47 [27] Benjamin-chung J, Abedin J, Berger D, Clark A, Jimenez V, Konagaya E, et al. Spillover effects on  
48 health outcomes in low- and middle-income countries: a systematic review. *International Journal of*  
49 *Epidemiology*. 2017;(September):1–26.  
50  
51  
52

- 1  
2  
3  
4 [28] Neema S, Atuyambe LM, Otolok-Tanga B, Twijukye C, Kambugu A, Thayer L, et al. Using a clinic  
5 based creativity initiative to reduce HIV related stigma at the Infectious Diseases Institute, Mulago  
6 National Referral Hospital, Uganda. *African health sciences*. 2012;12(2):231–239.  
7  
8  
9 [29] Smith Fawzi MC, Siril H, Li N, McAdam K, Ainebyona D, McAdam E, et al. Agents of Change:  
10 stepped wedge randomized controlled trial of the NAMWEZA intervention with people living with  
11 HIV and members of their social networks. *BMJ Global Health*. 2019;.  
12  
13  
14 [30] Kaaya S, Arthur G, Siril H, Lienert J, Li N, McAdam K, et al. Leadership of PLWH as Agents of  
15 Change: comparing HIV-related risk behavior of people attending ART clinics in Dar es Salaam with  
16 members of their social networks. In: 20th International AIDS Conference; 2014..  
17  
18  
19 [31] Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC medical research*  
20 *methodology*. 2006;6(1):54.  
21  
22  
23 [32] Zou J, Yamanaka Y, John M, Watt M, Ostermann J, Thielman N. Religion and HIV in Tanzania:  
24 influence of religious beliefs on HIV stigma, disclosure, and treatment attitudes. *BMC public health*.  
25 2009;9(1):75.  
26  
27  
28 [33] McCarthy B. Appreciative Inquiry: an alternative to behaviour management. *Dementia*.  
29 2017;16(2):249–253.  
30  
31  
32 [34] Siril HN, Kaaya SF, Fawzi MKS, Mtisi E, Somba M, Kilewo J, et al. CLINICAL outcomes and loss to  
33 follow-up among people living with HIV participating in the NAMWEZA intervention in Dar es  
34 Salaam, Tanzania: a prospective cohort study. *AIDS research and therapy*. 2017;14(1):18.  
35  
36  
37 [35] Carey MP, Schroder KE. Development and psychometric evaluation of the brief HIV Knowledge  
38 Questionnaire. *AIDS education and prevention*. 2002;14(2):172–182.  
39  
40  
41 [36] Haile BJ, Chambers JW, Garrison JL. Correlates of HIV knowledge and testing: Results of a 2003  
42 South African survey. *Journal of Black Studies*. 2007;38(2):194–208.  
43  
44  
45 [37] Aarø LE, Breivik K, Klepp KI, Kaaya S, Onya HE, Wubs A, et al. An HIV/AIDS knowledge scale for  
46 adolescents: item response theory analyses based on data from a study in South Africa and Tanzania.  
47 *Health education research*. 2011;26(2):212–224.  
48  
49  
50  
51  
52

- 1  
2  
3  
4 [38] Peltzer K, Matseke G, Mzolo T, Majaja M. Determinants of knowledge of HIV status in South Africa:  
5 results from a population-based HIV survey. *BMC public health*. 2009;9(1):174.  
6  
7 [39] Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. *Pharmaceutical Statistics: The*  
8 *Journal of Applied Statistics in the Pharmaceutical Industry*. 2009;8(1):50-61.  
9  
10 [40] Easterly W. How the millennium development goals are unfair to Africa. *World development*.  
11 2009;37(1):26–35.  
12  
13 [41] Barbarin OA, Khomo N. Indicators of economic status and social capital in South African townships:  
14 What do they reveal about the material and social conditions in families of poor children? *Childhood*.  
15 1997;4(2):193–222.  
16  
17 [42] Smith JA, McPherson M, Smith-Lovin L. Social distance in the United States: Sex, race, religion, age,  
18 and education homophily among confidants, 1985 to 2004. *American Sociological Review*. 2014  
19 Jun;79(3):432-56.  
20  
21 [43] Lesko CR, Edwards JK, Cole SR, Moore RD, Lau B. When to censor? *American journal of epidemi-*  
22 *ology*. 2017;187(3):623–632.  
23  
24 [44] Hernán MA. The hazards of hazard ratios. *Epidemiology (Cambridge, Mass.)*. 2010;21(1):13.  
25  
26 [45] VanderWeele T. *Explanation in causal inference: methods for mediation and interaction*. Oxford  
27 University Press; 2015.  
28  
29 [46] Hong, G., 2015. *Causality in a social world: Moderation, mediation and spill-over*. John Wiley & Sons.  
30  
31 [47] Forastiere L, Mealli F, Vanderweele TJ. Identification and Estimation of Causal Mechanisms in  
32 Clustered Encouragement Designs: Disentangling Bed Nets Using Bayesian Principal Stratification  
33 Identification and Estimation of Causal Mechanisms in Clustered Encouragement Designs : Disentan-  
34 gling Be. *Journal of the American Statistical Association*. 2016;111(514):510–525. Available from:  
35 <http://dx.doi.org/10.1080/01621459.2015.1125788>.  
36  
37 [48] Imai K, Jiang Z. Identification and sensitivity analysis of contagion effects in randomized placebo-  
38 controlled trials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2019.  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

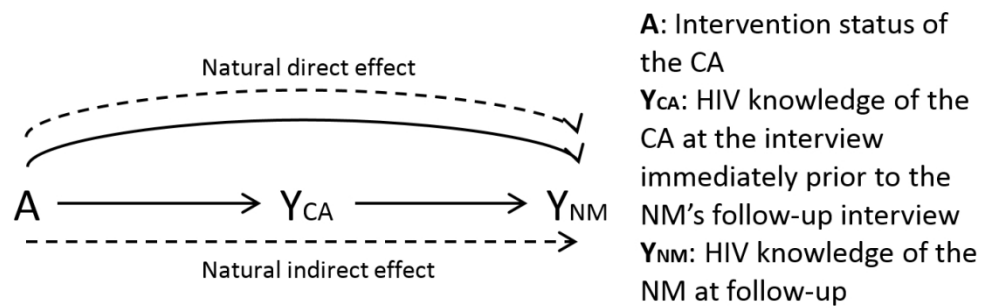
- 1  
2  
3 [49] Shalizi CR, Thomas AC. Homophily and contagion are generically confounded in observational social  
4 network studies. *SMR*. 2012;40(2):211–39.  
5  
6  
7 [50] De La Haye K, Robins G, Mohr P, Wilson C. Homophily and contagion as explanations for weight  
8 similarities among adolescent friends. *Journal of Adolescent Health*. 2011;49(4):421–427.  
9  
10  
11 [51] Vanderweele TJ, Hong G, Jones SM, Brown JL. Mediation and spillover effects in group-randomized  
12 trials: a case study of the 4Rs educational intervention. *Journal of the American Statistical Association*.  
13 2013;108(502):469–482. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23997375>.  
14  
15  
16 [52] Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal  
17 interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological*  
18 *methods*. 2013;18(2):137.  
19  
20  
21 [53] Leung MP. Treatment and spillover effects under network interference. *Review of Economics and*  
22 *Statistics*. 2016;p. 1–42.  
23  
24  
25 [54] Latkin CA, Davey-Rothwell MA, Knowlton AR, Alexander KA, Williams CT, Boodram B. Social  
26 network approaches to recruitment, HIV prevention, medical care, and medication adherence. *Journal*  
27 *of Acquired Immune Deiciency Syndromes*. 2013;63:1–9.  
28  
29  
30 [55] Galea S, Tracy M. Participation rates in epidemiologic studies. *Annals of epidemiology*.  
31 2007;17(9):643–653.  
32  
33  
34 [56] Berkman LF. Social support, social networks, social cohesion and health. *Social work in health care*.  
35 2000 Sep 26;31(2):3-14.  
36  
37  
38 [57] Langford CPH, Bowsher J, Maloney JP, Lillis PP. Social support: a conceptual analysis. *Journal of*  
39 *advanced nursing*. 1997;25(1):95–100.  
40  
41  
42 [58] Schneider JA, Zhou AN, Laumann EO. A new HIV prevention network approach: sociometric peer  
43 change agent selection. *Social science & medicine*. 2015;125:192–202.  
44  
45  
46 [59] He P, Eriksson F, Scheike TH, Zhang MJ. A proportional hazards regression model for the  
47 subdistribution with covariates-adjusted censoring weight for competing risks data. *Scandinavian*  
48 *Journal of Statistics*. 2016 Mar;43(1):103-22.  
49  
50  
51  
52



## 11. Figure Captions

Fig 1. Schematic of natural direct effect (NDE) and natural indirect effect (NIE). The NDE indicates the increase in NMs' HIV knowledge happens without a concomitant increase in their CA's HIV knowledge. The NIE, on the other hand, indicates that the increase in NMs' HIV knowledge is mediated by their CA's HIV knowledge increasing. Solid lines indicate paths of causality between variables. Dashed lines represent the line or lines composing the effect of interest.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



## STROBE (Strengthening The Reporting of OBServational Studies in Epidemiology) Checklist

A checklist of items that should be included in reports of observational studies. You must report the page number in your manuscript where you consider each of the items listed in this checklist. If you have not included this information, either revise your manuscript accordingly before submitting or note N/A.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).

Section and Item	Item No.	Recommendation	Reported on Page No.
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	
<b>Introduction</b>			
Background/Rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
<b>Methods</b>			
Study Design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed  <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	

Section and Item	Item No.	Recommendation	Reported on Page No.
Data Sources/ Measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study Size	10	Explain how the study size was arrived at	
Quantitative Variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	
Statistical Methods	12	(a) Describe all statistical methods, including those used to control for confounding	
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed  <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed  <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive Data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome Data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	

Section and Item	Item No.	Recommendation	Reported on Page No.
Main Results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other Analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key Results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
<b>Other Information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Once you have completed this checklist, please save a copy and upload it as part of your submission. DO NOT include this checklist as part of the main manuscript document. It must be uploaded as a separate file.**