

Assessment of Contamination and Misclassification Biases in a Randomized Controlled Trial of a Social Network Peer Education Intervention to Reduce HIV risk Behaviors Among Drug Users and Risk Partners in Philadelphia, PA and Chiang Mai, Thailand

Nicole Simmons¹ · Deborah Donnell² · San-san Ou² · David D. Celentano³ · Apinun Aramrattana⁴ · Annet Davis-Vogel⁵ · David Metzger⁵ · Carl Latkin⁶

Published online: 3 May 2015
© Springer Science+Business Media New York 2015

Abstract Controlled trials of HIV prevention and care interventions are susceptible to contamination. In a randomized controlled trial of a social network peer education intervention among people who inject drugs and their risk partners in Philadelphia, PA and Chiang Mai, Thailand, we tested a contamination measure based on recall of intervention terms. We assessed the recall of test, negative and positive control terms among intervention and control arm participants and compared the relative odds of recall of test versus negative control terms between study arms. The contamination measures showed good discriminant ability among participants in Chiang Mai. In Philadelphia there was no evidence of contamination and little evidence of diffusion. In Chiang Mai there was strong evidence of diffusion and contamination. Network structure and peer education in Chiang Mai likely led to contamination. Recall of intervention materials can be a useful method to detect contamination in experimental interventions.

Keywords Contamination · HIV · Prevention · Injection drug use · Social networks · Diffusion

Resumen Ensayos controlados de intervenciones de prevención y atención del VIH son susceptibles a la contaminación. En un ensayo controlado aleatorio de una red social intervención de educación inter pares entre personas que se inyectan drogas y sus socios de riesgo en Filadelfia, PA y Chiang Mai, Tailandia, probamos una medida contaminación basada en el recuerdo de los términos de intervención. Se evaluó el recuerdo de la prueba, las condiciones de control negativos y positivos entre los participantes de la intervención y del brazo de control y se compararon las probabilidades relativas (OR) de retirada de prueba vs. términos de control negativo entre los brazos del estudio. Las medidas de contaminación mostraron buena capacidad discriminante entre los participantes en Chiang Mai. En Filadelfia no había pruebas de contaminación y poca evidencia de la difusión. En Chiang Mai hubo una fuerte evidencia de la difusión y la contaminación. Estructura de la red y la educación entre pares en Chiang Mai probablemente llevaron a la contaminación. Llamada a una revisión de materiales de intervención puede ser un método útil para detectar la contaminación en las intervenciones experimentales.

✉ Carl Latkin
clatkin@jhsp.h.edu

¹ Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

² Statistical Center for HIV/AIDS Research and Prevention (SCHARP), Seattle, WA, USA

³ Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

⁴ Research Institute for Health Sciences, Chiang Mai University, Chiang Mai, Thailand

⁵ Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

⁶ Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, 737 Hampton House, Baltimore, MD 21205, USA

Introduction

Contamination bias in a randomized controlled trial (RCT) stems from exposure of participants in the control arm to materials in the intervention condition. Contamination generally biases the estimated treatment effect toward the null. Trials of educational prevention interventions are particularly vulnerable to such bias, as control arm

participants may be exposed to intervention arm messages directly, through the project inadvertently exposing controls to intervention materials or indirectly, through participants in the experimental condition interacting with those in control condition and providing them information, encouraging and modeling behavior change, or promoting new social norms. However, trials of such interventions rarely formally assess contamination [1]. Reported assessment methods include surveys of control group members on their exposure to intervention materials [1–3], and on their participation in intervention-like activities, particularly in trials of lifestyle interventions [4–6], collecting zip codes in geographically targeted online interventions [7] and comparisons of the results of individually randomized and cluster randomized trials.

Research on social diffusion—behavior change that travels through social networks—has become increasingly popular, as both infectious and non-infectious diseases have been shown to spread through social networks [8–10]. Using network members to educate their peers is a promising technique for advancing behavioral change [11, 12]. However, preventing and measuring contamination in such studies poses special challenges, as individuals may belong to multiple and shifting social networks [13, 14]. If the trial design utilizes social networks as a unit, links between social networks may be pathways for contamination. We illustrate these issues in a network oriented RCT to reduce HIV risk behaviors among injection drug users and their risk partners.

Relatively few studies have examined contamination between control and experimental conditions. In a review of studies of contamination among interventions with youth, Doyle and Hicky (2013) noted that contamination is rarely documented and that studies that choose a cluster randomized design rather than individual randomization may lose significant statistical power [16]. A review of over 150 health interventions by Keogh-Brown et al. (2007) reported that for high quality study designs there was a greater intervention effect for cluster randomized designs as compared to individual level interventions but concluded that for educational interventions there was only weak evidence of contamination bias [1].

Parent Study Description

This HIV Prevention Trials Network Trial 037, carried out in Philadelphia, PA, and Chiang Mai, Thailand, enrolled risk networks for acquisition of HIV infection. A detailed description of the study methodology and results is reported elsewhere [17]. Each network consisted of an index participant with a recent history of injection drug use and one or more social network members who (at baseline)

reported regularly injecting drugs and/or having sex with the index participant. Index participants listed and recruited members of their drug and sexual networks. All index participants were HIV-negative at baseline; network members could be either HIV-negative or HIV-positive. Enrolled networks were randomized to a group educational intervention—in which only index participants took part—or to a control condition consisting of voluntary HIV counseling and testing at each 6-month assessment. The educational intervention included development of a plan for encouraging risk reduction among participants' risk network members. The primary outcome was HIV seroconversion; changes in injection risk behaviors were secondary outcomes. Participants were followed up every 6 months for up to 24 months in Chiang Mai and 30 months in Philadelphia. HIV status was determined using standard laboratory assays and risk behaviors were assessed by an interviewer-administered questionnaire. All study protocols and procedures were approved by IRBs at Johns Hopkins University, University of Pennsylvania, Chiang Mai University, and the Thailand Ministry of Public Health. Voluntary written informed consent was provided by all participants.

Summary of Main Study Results

The study enrolled 414 networks with 1123 participants: 232 networks with 696 participants were enrolled in Philadelphia, and 182 networks with 427 participants were enrolled in Chiang Mai. Rates of injection risk behaviors declined dramatically between baseline and follow-up in both arms at both sites. At the Philadelphia site, intervention arm participants showed statistically significant reductions in a range of risky injection behaviors compared to the control arm. No significant differences between arms were observed at the Chiang Mai site.

Potential Explanations of differential site results

The study design was based in part on diffusion theory, which holds that change can be achieved by diffusion of information and behavior change through social networks [18]. We hypothesized that the disparate results in the two sites could be explained in whole or in part by differential level of diffusion – with greater diffusion occurring in Chiang Mai. Diffusion in Philadelphia was anticipated to be limited to the intervention arm network members, while in Chiang Mai intervention materials were anticipated to diffuse from the experimental condition to the control arm. Differential levels of contamination between the two sites was hypothesized since the injection drug user community in Philadelphia was relatively large, fluid, and geographically widespread. In northern Thailand some of the networks were

from small villages so it was likely that the controls and experimental participants may have known and interacted with each other. A drop-in center for participants in urban Chiang Mai (but not in Philadelphia) may also have facilitated contamination.

Methods

Selection of Diffusion Terms

The intervention included a number of specific phrases and terms designed to help participants remember the information they were taught. To assess diffusion of information, we asked participants whether they had discussed HIV risk reduction with others and if so, how many people, and tested whether terminology associated with the information were recognized by index and network participants at the 6-month follow-up visit. Participants were shown a list of terms and asked, “Which of these exact words or phrases have you heard before?” There were three groups of terms: test terms, negative control terms, and positive control terms. *Test terms* had been taught as part of the intervention sessions and were specific to the training program; these were similar but not identical at each site due to the differences in spoken languages. They included “SPEAKK,” an acronym used in the intervention as a mnemonic to help intervention index participants remember six specific communication skills they were taught, “injection risk ladder,” the mnemonic for the range of risk levels associated with various injection drug use behaviors and “Cleaning $1 \times 1 \times 1$,” which described a technique for cleaning needles and syringes. Index participants were exposed to some of these terms repeatedly during the intervention, while they may have heard others only once or twice.

Negative control terms were terms that were not used in the intervention or related to HIV prevention. We included these to assess tendencies of participants to state that they had heard the terms that might differ between the two arms and between index participants and network members. These were technical terms in common use by study staff at a given site (for instance, the acronym for the study statistical center, “SHARP”) but to which the participants should not have been exposed. *Positive control terms* were terms to which indexes and network members in both study arms may have been exposed; for example, the term “harm reduction,” which had been used in individual HIV counseling and testing sessions for both arms. This allowed us to estimate levels of recall of true exposures to intervention materials. In making the final selection of the terms for analysis, we excluded several terms that had been in use in drug education programs in the community, which participants might have been exposed to outside the study. For

this reason, “injection risk ladder” was excluded from analysis for the Chiang Mai site, and “cleaning $1 \times 1 \times 1$ ” was excluded from the Philadelphia site. The terms evaluated are listed in Table 1.

Analytic Methods

All analyses were site-specific. First, we tested the sensitivity of our analysis by examining our power to detect a difference in the odds of recalling positive control terms and the odds of recalling negative control terms. This comparison was made within each of the four subgroups—intervention arm index participants, intervention arm network members, control arm index participants and control arm network members. Second, we investigated whether the test terms were recalled by participants who were directly exposed to them during the intervention sessions by comparing the odds of index participants recalling a specific intervention term to their odds of recalling the negative control terms. Third, we investigated whether there was evidence of diffusion of test terms from index participants in the intervention arm to their own network members, by comparing the odds of network participants in the intervention arm recalling a test term to their odds of recalling the negative control terms. Fourth, we assessed whether there was evidence of diffusion of test terms to the control arm (i.e., contamination) by comparing the odds of control arm participants recalling a specific intervention term to their odds of they recalling the negative control terms. This test was performed separately for control arm index participants and control arm network members. For Philadelphia we also tested all control arm participants combined.

Among Chiang Mai participants, we assessed whether there was evidence that the extent of diffusion of terms to network members in the intervention arm and to index participants and network members in control arm was equivalent. We compared the relative odds of treatment members recalling each test term versus negative control terms was significantly different from the relative odds of control arm participants recalling test terms versus negative control terms. Finally, we explored whether the degree of exposure to the intervention, as measured by the number of intervention sessions attended by the index participant, affected recall of the intervention terms by indexes and network members in the intervention arm.

Analyses were carried out by logistic regression using generalized estimating equations (GEE) with unstructured correlation to account for the repeated observations for each participant. All regression models included indicator variables representing the type of term evaluated (intervention, positive control or negative control). Where appropriate, indicator variables were also incorporated to

Table 1 Terms tested in analysis of diffusion and contamination

	Term		
	Negative control	Positive control	Test
Philadelphia	EXPLORE Matrix method SCHARP	Harm reduction	Peer mentor SPEAKK Injection risk ladder Project FAST Ribbon game Sex risk ladder Freeze frame
Chiang Mai	EXPLORE Project FAST SCHARP data collection and analysis center	A friend who helps friends Harm reduction	6 communication skills (for a friend who helps friends) Ribbon game Cleaning 1 × 1 × 1 (1× water, 1× bleach, 1× water) Share the portion of heroin powder Time out role play

Negative control terms are terms to which no volunteers were directly exposed. Positive control terms were terms to which all volunteers should have been directly exposed. Test terms were those to which only index participants in the intervention arm were directly exposed

represent the participant group (treatment index, treatment network, control index, or control network). Statistical significance was evaluated at one-sided $p < 0.05$ level using SAS version 9.1. (SAS Institute, Cary, NC, 2006).

We also conducted a supplemental survey of a sample of 73 Chiang Mai participants to disentangle the possible routes of contamination at the Chiang Mai site—whether networks were unstable and/or broader than reported at enrollment, and whether participants expanded their networks to include individuals they met through the study. In a qualitative survey of a sub-sample of 24 of these participants, we asked about their friends' and their own drug use patterns, communication with friends about drug use and patterns of seeing friends. Supplemental analyses were conducted using Stata version 10.0. (Stata Corp., College Station Tex., 2007)

Results

Characteristics of the Study Participants

Of the 1123 participants who enrolled in the study, 954 (84.9 %) completed the 6 month visit and their recall of terminology was assessed. Participation was higher at the Thailand site (93.4 %) than the Philadelphia site (79.6 %). Within each site, participation was similar across the two study arms and between index participants and their network members (Table 2). Tables 3 and 4 show the demographic characteristics of the participants who responded to the questionnaire and relevant features of

their substance use behavior. In addition to the ethnic, linguistic and cultural differences between the two study sites, there were marked differences in demographics and drug use behaviors. Participants at both sites were overwhelmingly male, but the percentage was higher in Chiang Mai (82 %) than in Philadelphia (67 %). Participants tended to be older and better educated in Philadelphia than in Chiang Mai, but the Philadelphia participants were much less likely to be employed. Chiang Mai participants injected heroin and other drugs much less frequently than those in Philadelphia, but were much heavier users of alcohol. Non-injection drug use was common at both sites, but was more frequent in Philadelphia. The non-injection drugs of choice were different, with the primary drug used in Philadelphia being cocaine, and the primary drug in Thailand methamphetamine. Some participants at both sites reported smoking opiates and ingesting benzodiazepines. Network members at both sites were more likely than indexes to be female and reported less injection drug use. In Chiang Mai the network members also reported less alcohol and non-injection drug use than indexes.

Exposure to the treatment condition was higher among intervention index participants in Chiang Mai than in Philadelphia. In Philadelphia, 69 (75.8 %) of the 91 intervention index participants attended 5 or 6 of the 6 intervention sessions; 7 (7.7 %) Philadelphia indexes attended no intervention sessions at all. In Chiang Mai, by contrast, 79 (90.8 %) of the 87 intervention index participants attended 5 or 6 sessions, and only 1 (1.1 %) intervention index attended no sessions.

Table 2 Retention at 6 month visit

	Treatment Indexes		Treatment Network members		Control Indexes		Control Network members		Total	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Philadelphia										
Yes	91	81.3 %	172	76.8 %	98	81.7 %	194	80.8 %	555	79.7 %
No	21	18.7 %	52	23.2 %	22	18.3 %	46	19.2 %	141	20.3 %
Total	112		224		120		240		696	
Chiang Mai										
Yes	87	94.6 %	116	95.1 %	84	93.3 %	112	91.1 %	399	93.4 %
No	5	5.4 %	6	4.9 %	6	6.7 %	11	8.9 %	28	6.6 %
Total	92		122		90		123		427	
Overall										
Yes	178	87.3 %	288	83.2 %	182	86.7 %	306	84.3 %	954	85.0 %
No	26	12.7 %	58	16.8 %	28	13.3 %	57	15.7 %	169	15.0 %
Total	204		346		210		363		1123	

Number and percent of participants who completed the diffusion/contamination evaluation questionnaire, by site, type of participant and study arm

Conversations about HIV Prevention

At both sites, intervention index participants reported talking to significantly more people about ways to protect themselves against HIV infection than did control participants at the same sites. The mean number of persons talked with by Philadelphia index participants was 7 among the intervention arm versus 4 among the control arm ($p < 0.0001$); in Chiang Mai index participants spoke with 9 persons in intervention arm versus 6 in the control arm ($p < 0.0001$). HIV prevention was a significantly more frequent topic of conversation among all subgroups of Thai participants than among their U.S. counterparts.

Sensitivity of Recall Analysis

At both sites, our methodology distinguished between positive control terms and negative control terms. The contrast in rates of recall of positive versus negative control terms was much greater in Chiang Mai than in Philadelphia (Table 5). There were 134 (24.1 %) of 555 Philadelphia participants who recalled the positive control term “harm reduction.” By comparison, the negative control terms “SCHARP” and “Matrix Method” were each reported to be recognized by 41 (7.4 %) persons and the negative control term “EXPLORE” was recalled by 68 (12.3 %) of Philadelphia participants. In Chiang Mai, the same three negative control terms were recalled by 18–25 (4.5–6.3 %) of the 399 participants surveyed, compared to 280 (70.2 %) who recognized the positive control term “A friend who helps friends,” and 192 (48.1 %) who recognized the positive control term “harm reduction.”

Recall of Intervention Terms

In Philadelphia, treatment indexes demonstrated significantly greater recall of six of seven negative control terms. Rates of recognition of the intervention terms among the 91 treatment index participants surveyed ranged from 22 to 73 %. Only one term, “freeze frame,” was not clearly recognized by this group (Table 6). In Chiang Mai, treatment indexes clearly recalled all five of the test terms, at rates ranging from 64 to 89 %.

Diffusion of Intervention Terms

In Philadelphia, there was evidence of diffusion of only two of the seven test terms, “peer mentor” and “sex risk ladder,” from treatment arm indexes to their network members. The other five test terms assessed were not more likely to be recognized by treatment arm network members than were the negative control terms. Rates of recognition of the test terms among the 172 treatment network members surveyed ranged from 3 to 16 %. In Chiang Mai, four of the five test terms showed evidence of diffusion from treatment arm indexes to their network members. Rates of recognition of the test terms among the 116 treatment network members surveyed ranged from 7 to 41 %. The terms related to safer methods of drug use, “cleaning 1 × 1 × 1” and “share the portion of heroin powder,” were particularly likely to be recalled by treatment arm network members, with odds ratios of 12.4 (95 % CI 5.17–30.0) and 25.9 (95 % CI 10.7–62.9), respectively, compared to negative control terms (Table 7). Intervention session attendance by the index participants did not significantly influence their network members’ odds of recalling the intervention terms.

Table 3 Demographics and baseline risk behaviors of participants at Philadelphia, USA site

	Treatment Indexes <i>N</i> = 91	Treatment Networks <i>N</i> = 172	Control Indexes <i>N</i> = 98	Control Networks <i>N</i> = 194	Total <i>N</i> = 555
Gender					
Male	70 (77 %)	107 (62 %)	80 (82 %)	114 (59 %)	371 (67 %)
Female	21 (23 %)	65 (38 %)	18 (18 %)	80 (41 %)	184 (33 %)
Age					
18–30	16 (18 %)	20 (12 %)	20 (20 %)	31 (16 %)	87 (15 %)
31–40	25 (27 %)	49 (28 %)	24 (24 %)	59 (30 %)	157 (28 %)
40+	50 (55 %)	103 (60 %)	54 (55 %)	104 (54 %)	311 (56 %)
Education					
No or primary	0 (0 %)	2 (1 %)	3 (3 %)	4 (2 %)	9 (2 %)
Some secondary	70 (77 %)	131 (76 %)	75 (77 %)	150 (77 %)	426 (77 %)
Completed secondary	21 (23 %)	39 (23 %)	20 (20 %)	40 (21 %)	120 (22 %)
Employment					
Full time	5 (5 %)	16 (9 %)	9 (9 %)	15 (8 %)	45 (8 %)
Part time/occasional	11 (12 %)	14 (8 %)	13 (13 %)	19 (10 %)	57 (10 %)
Unemployed	75 (82 %)	142 (83 %)	76 (78 %)	160 (82 %)	453 (82 %)
Non-injection drug use					
Opiates (smoked)	31 (34 %)	70 (41 %)	27 (28 %)	66 (34 %)	194 (35 %)
Benzodiazepines	45 (49 %)	76 (44 %)	54 (55 %)	106 (55 %)	281 (51 %)
Crack/Cocaine (smoke)	62 (68 %)	92 (53 %)	48 (49 %)	99 (51 %)	301 (54 %)
Cocaine (snort or sniff)	14 (15 %)	32 (19 %)	16 (16 %)	33 (17 %)	95 (17 %)
Amphetamines (smoke)	0 (0 %)	3 (2 %)	1 (1 %)	1 (< 1 %)	5 (< 1 %)
Alcohol use					
None	42 (46 %)	55 (32 %)	41 (42 %)	79 (41 %)	217 (39 %)
Yes, but never drunk	14 (15 %)	53 (31 %)	34 (35 %)	48 (25 %)	149 (27 %)
Sometimes/always drunk	35 (38 %)	64 (37 %)	23 (23 %)	66 (34 %)	188 (34 %)
Days injected in last month					
0	1 (1 %)	21 (12 %)	2 (2 %)	30 (15 %)	54 (10 %)
1–5	0 (0 %)	19 (11 %)	2 (2 %)	19 (10 %)	40 (7 %)
6–29	36 (40 %)	51 (30 %)	30 (31 %)	47 (24 %)	164 (30 %)
Everyday	54 (59 %)	81 (47 %)	64 (65 %)	98 (51 %)	297 (54 %)

Contamination of Control Arm

In Philadelphia, there was some evidence that one of the seven tested intervention terms, “peer mentor,” had diffused from the treatment arm to the control arm. For control arm indexes, the odds ratio of recalling this term compared to the negative control terms was 2.09 (95 % CI 1.08–4.04). However, among control arm network participants, the odds were non-significant: (1.30 (95 % CI 0.74–2.31)). None of the other six intervention terms was significantly more likely to be recognized by participants in the control arm (Table 6). In Chiang Mai, there was strong evidence that four of the five intervention terms had diffused from the intervention arm to the control arm; from 12 to 44 % of the 196 control arm members recalled these terms. Only one term, “ribbon game,” was not significantly more likely to

be recalled by control arm participants than the negative control terms. The evidence of diffusion to the control arm was particularly strong for the two terms related to safer methods of drug use (Table 7). Terms appeared just as likely to have diffused to control arm participants as to treatment arm network members; there was no significant difference in the relative odds of recognition of any of the five intervention terms, compared to negative control terms, between treatment network members and either of the two control arm subgroups (data not shown).

Composition of Participants’ “Real” Social Networks

In the Chiang Mai supplemental survey (*N* = 73), participants were asked how many of the friends they had

Table 4 Demographics and baseline risk behaviors of participants at Chiang Mai, Thailand site

	Treatment Indexes <i>N</i> = 87	Treatment Networks <i>N</i> = 116	Control Indexes <i>N</i> = 84	Control Networks <i>N</i> = 112	Total <i>N</i> = 399
Gender					
Male	84 (97 %)	86 (74 %)	81 (96 %)	78 (70 %)	329 (82 %)
Female	3 (3 %)	30 (26 %)	3 (4 %)	34 (30 %)	70 (18 %)
Age					
18–30	46 (52 %)	59 (51 %)	43 (51 %)	67 (60 %)	215 (54 %)
31–40	23 (26 %)	24 (21 %)	21 (25 %)	23 (21 %)	91 (23 %)
40+	18 (21 %)	33 (28 %)	20 (24 %)	22 (20 %)	93 (23 %)
Education					
No or primary	55 (63 %)	85 (73 %)	52 (62 %)	74 (66 %)	266 (67 %)
Some secondary	29 (33 %)	26 (22 %)	26 (31 %)	31 (28 %)	112 (28 %)
Completed secondary	3 (3 %)	5 (4 %)	6 (7 %)	7 (6 %)	21 (5 %)
Employment					
Full time	58 (67 %)	78 (67 %)	46 (55 %)	71 (63 %)	253 (63 %)
Part time/occasional	20 (23 %)	30 (26 %)	30 (36 %)	25 (22 %)	105 (26 %)
Unemployed	9 (10 %)	8 (7 %)	8 (10 %)	16 (14 %)	41 (10 %)
Non-injection drug use					
Opiates (smoked)	19 (22 %)	23 (20 %)	14 (17 %)	13 (12 %)	69 (17 %)
Benzodiazepines	10 (11 %)	8 (7 %)	12 (14 %)	7 (6 %)	37 (9 %)
Crack/Cocaine (smoke)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
Cocaine (snort or sniff)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
Amphetamines (smoke)	35 (40 %)	42 (36 %)	41 (49 %)	39 (35 %)	157 (39 %)
Alcohol use					
None	2 (2 %)	23 (20 %)	8 (10 %)	23 (21 %)	56 (14 %)
Yes, but never drunk	27 (31 %)	27 (23 %)	26 (31 %)	28 (25 %)	108 (27 %)
Sometimes/always drunk	58 (67 %)	66 (57 %)	50 (60 %)	61 (54 %)	235 (59 %)
Days injected in last month 0					
1–5	10 (11 %)	39 (34 %)	13 (15 %)	49 (44 %)	111 (28 %)
6–29	35 (40 %)	41 (35 %)	35 (42 %)	37 (33 %)	148 (37 %)
Everyday	38 (44 %)	26 (22 %)	32 (38 %)	18 (16 %)	114 (29 %)
Everyday	4 (5 %)	10 (9 %)	4 (5 %)	8 (7 %)	26 (7 %)

Table 5 Assessment of sensitivity of recall analysis: relative odds of recall of positive versus negative control terms, by site and subgroup

	Treatment arm indexes			Treatment arm network members			Control arm indexes			Control arm network members		
	<i>N</i>	OR	<i>p</i> value	<i>N</i>	OR	<i>p</i> value	<i>N</i>	OR	<i>p</i> value	<i>N</i>	OR	<i>p</i> value
Philadelphia	91	4.80	<0.0001	172	2.15	0.002	98	4.39	<0.0001	194	3.17	<0.0001
Chiang Mai	87	35.95	<0.0001	116	33.19	<0.0001	84	16.27	<0.0001	112	19.26	<0.0001

Results of unadjusted logistic regression with GEE unstructured correlation, conducted separately for each site and subgroup. *P*-values are one-sided

known before enrolling in the study enrolled in the study. The size of the “real” social networks among study participants reported in this survey ranged from 0 to 30, with a median size of 4. The number of other members of each participant’s enrolled network ranged from 1 (the minimum number required by the study protocol) to 3, with a median size of 1. Fifty-five (75 %) of the 73 participants

surveyed, including 26 (76 %) of the 34 index participants surveyed reported that—at the study start—they had more friends in the study than were co-enrolled in their network. Twenty-five (34 %) of the participants said they had made new friends through the program. The number reporting making new friends included 16 (30 %) of the 54 participants who were either in the control arm or were

Table 6 Recognition of test terms, compared to negative control terms, at the Philadelphia site, by subgroup

Term	Treatment arm indexes (<i>N</i> = 91)		Treatment arm network members (<i>N</i> = 172)		Control arm indexes (<i>N</i> = 98)		Control arm network members (<i>N</i> = 194)	
	OR	<i>p</i> value	OR	<i>p</i> value	OR	<i>p</i> value	OR	<i>p</i> value
Peer mentor	10.34	<0.0001	2.94	<0.0001	2.09	0.014	1.30	0.18
SPEAKK	5.76	<0.0001	0.47	0.96	0.88	0.62	0.61	0.87
Injection risk ladder	5.50	<0.0001	1.38	0.14	1.13	0.37	0.99	0.51
Project FAST	2.27	<0.0001	0.85	0.66	1.75	0.08	1.03	0.46
Ribbon game	4.34	<0.0001	0.38	0.97	0.37	0.96	0.45	0.97
Sex risk ladder	7.20	<0.0001	1.82	0.026	1.36	0.20	1.26	0.24
Freeze frame	1.08	0.36	1.19	0.27	0.56	0.89	1.40	0.08

Results of unadjusted logistic regression with GEE unstructured correlation, conducted separately for each subgroup. *p* values are one-sided

Table 7 Recognition of test terms, compared to negative control terms, at the Chiang Mai Site, by subgroup

Term	Treatment arm indexes (<i>N</i> = 87)		Treatment arm network members (<i>N</i> = 116)		Control arm indexes (<i>N</i> = 84)		Control arm network members (<i>N</i> = 112)	
	OR	<i>p</i> value	OR	<i>p</i> value	OR	<i>p</i> value	OR	<i>p</i> value
6 communication skills	53.09	<0.0001	4.93	0.0001	3.78	<0.0001	3.24	0.0071
Ribbon game	23.39	<0.0001	2.16	0.067	0.89	0.64	0.42	0.88
Cleaning 1 × 1 × 1	53.09	<0.0001	12.44	<0.0001	6.62	<0.0001	7.42	<0.0001
Share the portion of heroin powder	34.03	<0.0001	25.93	<0.0001	15.32	<0.0001	21.97	<0.0001
Time out role play	12.26	<0.0001	4.05	0.0011	3.17	0.0011	2.91	0.0006

Results of unadjusted logistic regression with GEE unstructured correlation, conducted separately for each subgroup. *p* values are one-sided

network members in the treatment arm, potentially leading to contamination. All 25 said they had met at least some of their new friends through sources other than the intervention group, such as the study drop-in center, through other friends in the study, or out in the community. Some participants also reported no longer belonging to the social networks of the friends with whom they enrolled in the study. Twenty-seven (38 %) of the 71 participants who said a friend told them about the study said they had not seen that friend in 3 months. Ten (23 %) of the 43 who encouraged someone else to join the study said they had not seen that friend in 3 months.

Discussion

Very few RCTs of HIV prevention interventions have directly examined contamination. One study by Lang and colleagues found little evidence of contamination but reported that their measures of contamination were “relatively crude” [19]. The present study include a range of measures of contamination. At the Thailand site, our

analysis had sufficient discriminative power to determine that terminology and concepts taught to participants in the intervention arm of the study diffused, as intended, to members of their social networks. This terminology likely diffused through the conversations intervention indexes had about HIV prevention, which were significantly more common among the intervention arm. However, the people they talked to about HIV risk reduction likely included not only those co-enrolled with them in the treatment arm of the study, but also others enrolled in the control arm. The recognition rates for the terminology and concepts taught in the intervention were indistinguishable between the intervention network members and participants enrolled in the control arm of the study. These results supported the hypothesis that there was evidence of contamination in Chiang Mai.

While all five intervention terms tested were highly likely to be recalled by intervention indexes, with odds ratios for recall greater than 10, only two of the terms showed very strong recall by the other three groups: “Cleaning 1 × 1 × 1” and “Share the portion of heroin powder,” for which odds ratios of recollection ranged from

6.6 to 25.9. Among the terms evaluated for recall, these were the terms most directly related to safer drug use techniques. The other three terms, “6 communication skills,” “Ribbon Game,” and “Time Out Role Play,” were related more closely to the intervention process, and indexes might not find these as relevant to communicate to their network members. The communication of these terms may serve as one mediator of the dramatic and approximately equal reduction of drug use risk behaviors in both study arms at the Thai site.

At the Philadelphia site, despite statistically significant results demonstrating the effectiveness of the intervention in reducing drug use risk behaviors, there was little evidence of diffusion of intervention terms. Therefore, our analysis may not have had sufficient discriminative power to adequately test whether terminology taught in the intervention arm diffused to others. Treatment arm network participants and control arm participants reported that they recognized few terms, even terms to which they had been repeatedly exposed. Regarding contamination, there was some evidence that the term “peer mentor” may have diffused to the control arm. However, this term was recognized above background levels solely by control indexes.

It is unclear why participants outside the intervention arm were so much less likely to report recalling any of the tested terms at the US site than at the Thai site. It is possible the US index participants, as compared to their Thai counterparts, were less likely to describe verbatim the content of the intervention to their network members and adapted the risk reduction messages to their own communication style. All categories of Philadelphia participants were significantly less likely than those in Thailand to report talking about HIV prevention. While nearly all (98 %) of Thai treatment indexes reported talking about HIV prevention, only 75 % of Philadelphia treatment indexes reported conversations about protecting oneself from HIV. It may be that Philadelphia injectors are more likely to model behavior change for their peers rather than using words to describe it. Not having repeated the terms from the intervention, they may be less likely to recall them.

The lack of recollection may have been due to the lower level of attendance in the intervention in the US compared to Thailand. However, there was no association within the Philadelphia cohort between indexes’ session attendance and their network members’ recognition of intervention terms. This lack of an association may be also due to lack on power to detected differences by number of sessions the indexes attended as well differences in frequency of interactions with their risk network members.

Limitations in the study design might also have contributed to the differences between sites. We initially attempted to test a similar range of terms at each site, but later excluded certain terms from the analysis due to their

use in the general community. Those terms that were most concretely linked to safe drug use techniques showed the greatest penetration in Chiang Mai. It might have been more effective to test such terms in Philadelphia. Unfortunately, those terms were the ones that overlapped with terms use in other community programs, and hence we were not able to test them. It would have been useful to have conducted a parallel sub-study in Philadelphia to investigate in more detail how and to whom index participants communicated. We would like to know whether they used terminology from the intervention directly, rephrased the concepts, or simply modeled the desired behaviors.

Our analysis shows both the promise and the limitations of using the recall technique to detect diffusion of terminology and contamination of educational messages between arms in a controlled trial of an educational intervention. Several lessons can be drawn. Such an analysis will work best if it is possible to use terminology in the educational intervention that are fully distinct from terms to which participants are exposed from other sources. Investigators should conduct a careful survey of other educational activities in the surrounding community when designing both the intervention and the diffusion assessment, to ensure that the same terms are not being used elsewhere. The diffusion assessment should be developed and piloted as part of the development of the intervention to ensure the terminology selected is memorable. Multiple positive and negative control terms should also be selected and pilot tested.

A measure of diffusion of an educational term or message can also be used to assess whether the educational message was a mediating factor in behavior change. Measurement of contamination can be used to adjust the estimated results of controlled trials. If each participant’s exposure level to the intervention (whether through contamination or compliance) is known, contamination by an educational intervention can be treated in the same manner as non-compliance with a clinical intervention, and the statistical methods that have been developed to handle non-compliance used to address it. These include G-estimation, structural models and related methods reviewed by Robins [20], calculation of the Complier Average Causal Effect (CACE) [1, 21, 22] using an instrumental variables approach to restrict the analysis to participants not exposed to contamination and a comparable group in the intervention arm [23], and using propensity scores [24] to estimate propensity of exposure to treatment as in an observational study, but with inclusion of treatment assignment as one of the covariates. Unfortunately, all these approaches assume contamination occurs among participants independently and applications to network-based studies have not been fully developed.

Our analysis also demonstrates the importance of considering the possibility of overlapping and shifting social networks when designing controlled studies of network-based educational interventions. Care should be taken to ensure that networks exposed to the intervention are isolated from those in the control condition and that the study does not inadvertently provide an opportunity for those networks to overlap. Alternatively, interventions may train indexes in the experimental condition to have HIV-related conversations with network members who are also in the experimental condition. Studies should consider the network structure and stability and whether participants are likely to promote behavior change with small or large number of their network members. Cluster-randomization is one technique that might be considered to prevent contamination in social network-based studies.

Acknowledgments This study was supported by the HIV Prevention Trials Network (HPTN) and sponsored by the National Institute of Allergy and Infectious Diseases, National Institute of Child Health and Human Development, National Institute on Drug Abuse, National Institute of Mental Health, and Office of AIDS Research, of the National Institutes of Health, U.S. Department of Health and Human Services, through cooperative agreement U01-AI-46749 with Family Health International, U01-AI-46702 with Fred Hutchinson Cancer Research Center, U01-AI-47984 with Johns Hopkins University, and U01-AI-48014 with the University of Pennsylvania. The authors gratefully acknowledge the contributions of Thira Sirisanthana MD, Tساناي Vongchak RN MPH, Namtip Srirak PhD, Antika Wongthane, Kanokporn Wiboonnatakul, Lara Siree Johnson, and Chatsuda Auching.

References

- Keogh-Brown MR, Bachmann MO, Shepstone L, et al. Contamination in trials of educational interventions. *Health Technol Assess*. 2007;11(43):1–128.
- Tilgren P, Dignan M, Michielutte R. Assessment of contamination in a trial of community-based cancer education. *Am J Health Behav*. 1998;24(4):292–7.
- Ross MW, Chatterjee NS, Leonard L. A community level syphilis prevention programme: outcome data from a controlled trial. *Sex Transm Inf*. 2004;80:100–4.
- Moore GF, Williams A, Moore L, Murphy S. An exploratory cluster randomised trial of a university halls of residence based social norms marketing campaign to reduce alcohol consumption among 1st year students. *Subst Abuse Treat Prev Policy*. 2013; 8:15.
- Stewart-Brown S, Patterson J, Mockford C, Barlow J, Klimes I, Pyper C. Impact of a general practice based group parenting programme: quantitative and qualitative results from a controlled trial at 12 months. *Arch Dis Child*. 2004;89(6):519–25.
- Courneya KS, Friedenreich CM, Sela RA, Quinney HA, Rhodes RE, Handman M. The group psychotherapy and home-based physical exercise (group-hope) trial in cancer survivors: physical fitness and quality of life outcomes. *Psychooncology*. 2003;12(4): 357–74.
- van Sluijs EMF, van Poppel MNM, Twisk JWR, van Mechelen W. Physical activity measurements affected participants' behavior in a randomized controlled trial. *J Clin Epidemiol*. 2006;59(4): 404–11.
- Jones RB, Goldsmith L, Hewson P, Williams CJ. Recruitment to online therapies for depression: pilot cluster randomized controlled trial. *J Med Internet Res*. 2013;15(3):e45.
- Doherty IA, Padian NS, Marlow C, Aral SO. Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. *J Infect Dis*. 2005;191(Suppl 1): S42–54.
- Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *N Engl J Med*. 2007;357(4):370–9.
- Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *N Engl J Med*. 2008;358(21):2249–58.
- Latkin CA, Sherman S, Knowlton A. HIV prevention among drug users: outcome of a network-oriented peer outreach intervention. *Health Psychol*. 2003;22(4):332–9.
- Latkin CA, Mandell W, Vlahov D, Oziemkowska M, Celentano DD. The long-term outcome of a personal network-oriented HIV prevention intervention for injection drug users: the SAFE Study. *Am J Community Psychol*. 1996;24(3):341–64.
- Liljeros F, Edling CR, Nunes Amaral LA. Sexual networks: implications for the transmission of sexually transmitted infections. *Microbes Infect*. 2003;5(2):189–96.
- Palla G, Barabasi A-L, Vicsek T. Quantifying social group evolution. *Nature*. 2007;446:664–7.
- Doyle O, Hickey C. The challenges of contamination in evaluations of childhood interventions. *Evaluation*. 2013;19(2):183–96.
- Latkin C, Donnell D, Metzger D, et al. The efficacy of a network intervention to reduce HIV risk behaviors among drug users and risk partners in Chiang Mai, Thailand and Philadelphia, USA. *Soc Sci Med*. 2009;68:740–8.
- Rogers E. *Diffusion of innovations*. 5th ed. New York: Free Press; 2003.
- Lang DL, DiClemente RJ, Hardin JW, et al. Threats of cross-contamination on effects of a sexual risk reduction intervention: fact or fiction. *Prev Sci*. 2009;10(3):270–5.
- Robins JM. Correction for non-compliance in equivalence trials. *Stat Med*. 1998;17:269–302.
- Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Stat Med*. 1997;16:1017–29.
- Stuart EA, Perry DF, Le H-N, Ialongo NS. Estimating intervention effects of prevention programs: accounting for noncompliance. *Prev Sci*. 2008;9(4):288–98.
- Sussman JB, Hayward RA. An IV for the RCT: using instrumental variables to adjust for treatment contamination in randomised controlled trials. *BMJ*. 2010;340:c2073.
- Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41–55.