2018

Assessing Individual and Disseminated Effects in Network-Randomized Studies

Ashley L. Buchanan
University of Rhode Island, buchanan@uri.edu

H. Vermund

Samuel R. Friedman

Donna Spiegelman

Citation/Publisher Attribution

Ashley L Buchanan, Sten H Vermund, Samuel R Friedman, Donna Spiegelman; Assessing Individual and Disseminated Effects in Network-Randomized Studies, *American Journal of Epidemiology*, , kwy149, https://doi.org/10.1093/aje/kwy149
Available at: https://doi.org/10.1093/aje/kwy149

Assessing Individual and Disseminated Effects in Network-Randomized Studies

Assessing Individual and Disseminated Effects in Network-Randomized Studies

Ashley L. Buchanan, Sten H. Vermund, Samuel R. Friedman, and Donna Spiegelman

Correspondence to Dr. Ashley Buchanan, Department of Pharmacy Practice, College of Pharmacy, University of Rhode Island, 7 Greenhouse Road, Kingston, RI 02881 (email: Buchanan@uri.edu; phone: 401-874-4739; fax: 401-874-2717)

Author affiliations: Department of Pharmacy Practice, College of Pharmacy, University of Rhode Island, Kingston, Rhode Island (Ashley L. Buchanan); Vanderbilt Institute for Global Health and Department of Pediatrics[2], Vanderbilt University School of Medicine, New Haven, Connecticut (Sten H. Vermund); Institute for Infectious Disease Research, National Development and Research Institutes, Inc., New York, New York (Samuel R. Friedman); Departments of Epidemiology, Biostatistics, Nutrition, and Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA (Donna Spiegelman)

Funding: ALB and DS were supported by the National Institutes of Health (NIH) grants 5R01AI112339-02 and 1DP1ES025459 (DP1). SHV was supported by NIH grants UM1AI068619, 5R01AI112339, and P30AI110527. SRF was supported by NIH grants DP1DA034989 and P30DA011041. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of interest: None declared.

Running Head: Network-Randomized Studies with Dissemination

Abbreviations: Confidence interval (CI), Generalized estimating equations (GEEs), HIV Prevention Trials Network (HPTN), Human Immunodeficiency Virus (HIV), Risk/Rate difference (RD), Risk/Rate ratio (RR)

ABSTRACT

Implementation trials often involve clustering via risk networks, where only some participants directly received the intervention. The individual effect is that among directly treated persons beyond being in an intervention network; the disseminated effect is that among persons engaged with those directly treated. We employ a causal inference framework and discuss assumptions and estimators for individual and disseminated effects and apply them to HIV Prevention Trials Network 037. HIV Prevention Trials Network 037 was a Phase III, network-level, randomized controlled HIV prevention trial conducted in the US and Thailand from 2002 to 2006 that recruited persons who injected drugs, who received either intervention or control, and their risk network members, who received no direct intervention. Combining individual and disseminated, a 35% composite rate reduction was observed in the adjusted model (95% confidence interval = 0.47, 0.90). Methodology is now available to estimate the full set of these effects enhancing knowledge gained from network-randomized trials. Although the overall effect gains validity from network randomization, we show that it will, in general, be less than the composite effect. Additionally, if only index participants benefit from the intervention, as the network size increases, the overall effect tends to the null, an unfortunate and misleading conclusion.

Keywords: Causal inference; Cluster-randomized trials; Disseminated/Indirect effects; Drug use/abuse; HIV/AIDS; Implementation Science; Individual/Direct effects; Interference

Implementation science studies how best to translate and scale-up research evidence into practice. These studies often involve a natural clustering by social network, facility or community. In such network-randomized trials, only some members of networks randomized to the intervention directly receive the intervention. Individual and disseminated effects can be targets of inference. The disseminated or indirect effect is the impact of the intervention on the network members who were not directly exposed in the intervention networks. The individual or direct effect is the impact of the intervention on those who directly received the intervention, the index participants in intervention networks, beyond being an intervention network member. The composite effect is the effect of the intervention on index participants compared to network control members; that is, the maximal attainable benefit of the intervention. For example, a health care professional may educate an index participant, also known as the *ego*, who then in turn may educate or otherwise influence members of his or her risk network, also known as the *alter-egos*, to modify their risky practices. In this setting, there is interest in the intervention effect on those who were directly educated as well as those sharing risk networks with the index.

The terms individual and disseminated are used in this paper to avoid confusion with these terms used in the mediation literature, where the direct and indirect effects are terms used to describe parameters addressing different scientific questions (1). We present a summary of the vernacular from relevant literature on the present topic and provide our recommended terms in Table 1. Compared to previous terminology, our terms are more agnostic to the relative magnitudes and desirability of dissemination.

In causal inference methodology, a fundamental assumption of much work is the **s**table **u**nit **t**reatment **v**alue **a**ssumption, that is, SUTVA (2), which includes an assumption of no dissemination, or interference, between individuals. No dissemination requires that the potential outcomes of one individual are unaffected by the intervention assignment of other individuals. In this paper, the primary research interest is precisely in relaxing the no dissemination assumption of the stable unit treatment value assumption and quantifying disseminated effects.

Earlier work on methods for assessing dissemination assumed two-stage randomization, where networks were first randomized to an intervention allocation strategy and, then, within a network, individuals were randomized according to their network's allocation strategy (3, 4). Estimators of individual and disseminated effects were motivated by vaccine studies, where herd immunity is a good example of a disseminated effect (5, 6). Permutation-based variance estimators were developed for these doubly-randomized designs (4). When a study is not doubly randomized, these estimators of individual and disseminated effects are no longer valid because of the potential for bias due to confounding at either the network-, individual- level or both. Tchetgen Tchetgen and VanderWeele (7) proposed inverse probability weighted estimators of individual and disseminated effects for studies where randomization is not required at the individual or group level. This approach was applied to an individually-randomized study of cholera vaccine, where individuals were clustered by groups of households. (8). This approach was then applied to a study design with more than one index member and unequal treatment probabilities in the second stage (9).

Using multivariable outcome models rather than inverse probability weighted models, we develop alternative estimators of individual, disseminated and composite causal effects for a setting with randomization only at the network level and one index per network, a study design frequently utilized in drug abuse and addiction research (10-12), and provide methods for asymptotic inference. We discuss the causal inference framework and assumptions for this setting. We prove some general results of interest in this setting that demonstrate the utility of the methods proposed. We apply these methods to a network-randomized trial in the HIV Prevention Trials Network (HPTN)(13-15) to obtain estimates of the individual, disseminated, composite, and overall intervention effects. Lastly, we discuss some limitations of this approach and identify future methodological directions for causal inference in network-randomized studies.

METHODS

Assumptions and Notation

The sufficient conditions for valid estimation of causal effects have been previously described (16) . We assume no dissemination between networks. Because the networks are randomized to the intervention, on average, exchangeability at the network level holds. Networks randomized to the intervention will be, on average, comparable to networks randomized to the control. There is an additional exchangeability assumption that allows for valid estimation of all the parameters of interest in this setting. Within each network, conditional on a set of measured covariates, the potential outcomes of the voluntary index participants and non-index network members are the same as, on average, that would be expected if any other network member volunteered to be the index, whether the network was randomized to the intervention or not. We call this "conditional index exchangeability".

Let K be the total number of networks, $k = 1, \dots, K$, with $i = 1, \dots, n_k$ participants in network k, where each participant i has $j = 1, \dots, m_{ki}$ visits and $\sum_{k=1}^{K} n_k = N$ is the total sample size of the study. Let $\boldsymbol{Z}_{ki}$ be a vector of measured baseline covariates for participant i in network k. Define Y_{kij} as the outcome for the i^{th} participant in the k^{th} network at the j^{th} visit. Let X_k be the network-level intervention in network k assigned at the start of the trial and define an indicator R_{ki} for the individual-level index status, where $R_{ki} = 1$ if participant i in network k is an index and $R_{ki} = 0$ otherwise. For example, investigators assign HIV risk networks either to train their index member to be peer educator or standard of care with HIV counseling and testing. In network k, we have an n_k-vector of observed index status indicators $\boldsymbol{R}_k = \left(R_{k1}, R_{k2}, \dots, R_{kn_k}\right)$, constrained in this paper to $\sum_{i=1}^{n_k} R_{ki} = 1$ and n_k-vectors of baseline covariates $\boldsymbol{Z}_k = \left(Z_{k1}, Z_{k2}, \dots, Z_{kn_k}\right)$. Each participant has potential outcomes $\boldsymbol{Y}_{kij}(\boldsymbol{r}, x)$, which correspond to the $2 \times n_k$ vector of potential outcomes for individual i in network k at time j under the index status indicator vector $\boldsymbol{R}_k = \boldsymbol{r}$ and intervention assignment $X_k = x$. Because an individual's potential outcomes depend on the network-level intervention, dissemination is possible in this setting. A contrast between any two of these potential outcomes is a measure of a causal effect. For example, a representation of the individual causal effect in an intervention network where the second participant was the index compared to when the last participant was the index is $E\{\boldsymbol{Y}_{kij}[(0,1,\dots 0),1]\} - E\{\boldsymbol{Y}_{kij}[(0,1,\dots 1),1]\}$.

Because the number of index members in network k is fixed to 1 by design, there are $J_k = \binom{n_k}{1} = n_k$ possible configurations of index participants in each network k. In general, each network member has $2 \times n_k$ potential outcomes with n_k corresponding to intervention and n_k corresponding to control**.** With these many potential outcomes within networks, it is difficult to choose the causal effects of interest. Conditional on baseline covariates, we assume that there is exchangeability between the $2 \times n_k$ possible configurations when there is one index participant in network k; that is, $Y_{kij}(\boldsymbol{r}, x) \perp \boldsymbol{R}_k | \boldsymbol{Z}_k$. Baseline covariates are sufficient to control for confounding of the effect of self-selected index status on the outcome. Because there is only one index per network, we can validly denote the potential outcomes by $R_{ki} = r$ and $X_k = x$. Under these circumstances, the number of potential outcomes for each participant, $Y_{kij}(r, x)$, is reduced to four; that is, two for each of the two possible network-level intervention assignments, $Y_{kij}(1,1)$, $Y_{kij}(1,0)$, $Y_{kij}(0,1)$, and $Y_{kij}(0,0)$. For example, let $Y_{kij}(1,1)$ be the potential outcome of participant i at visit j in network k, if, possibly contrary to fact, this participant was an index member in a network randomized to the intervention.

Causal Framework and Estimands

The individual effect is defined as the effect of the intervention among index members in intervention networks beyond being in an intervention network (Figure

1), that is, Risk/Rate Difference^{I} (RD^{I}) $= E\left[Y_{kij}(1,1)\right] - E\left[Y_{kij}(0,1)\right]$. The disseminated effect is defined as the intervention effect among the non-index network members, that is, $RD^{D} = E\left[Y_{kij}(0,1)\right] - E[Y_{kij}(0,0)]$. Composite and overall effects combine the individual and disseminated effects in two different ways. The composite effect is the sum of the disseminated and individual effects, that is $RD^{Comp} = E\left[Y_{kij}(1,1)\right] - E\left[Y_{kij}(0,1)\right] + \left\{E\left[Y_{kij}(0,1)\right] - E\left[Y_{kij}(0,0)\right]\right\} = E\left[Y_{kij}(1,1)\right] - E\left[Y_{kij}(0,0)\right]$. The composite effect is the maximum possible effect of the intervention; that is, the effect of being an index member in an intervention network compared to a network member in a control network.

The overall effect is the average effect among all intervention network members compared to all control network members. In Web Appendix 1, we derive the parametric relationship between the individual and disseminated effects and the overall effect, and show that when network sizes vary, $RD^{Overall} = E\left[Y_{kij}(\cdot,1)\right] - E\left[Y_{kij}(\cdot,0)\right] = RD^{D} + RD^{I} \times E_{R}[R_{ki}] = RD^{D} + RD^{I} \times \frac{K}{N}$, which will always be smaller than $RD^{Comp} = RD^{D} + RD^{I}$ as long as RD^{I} and RD^{D} have the same sign, as would typically be the case. When network sizes are constant with $n_k = n$ for all k, $RD^{Overall} = RD^{D} + RD^{I} \times \frac{1}{n}$. When the sign differs, the overall effect will be smaller than the composite only in certain cases. For example, if the average network size is 3, $RD^{I} = 1$, and $RD^{D} = -3$, then $RD^{Comp} = -2$, which is smaller than $RD^{Overall} = -2.67$. It is typically not expected for individual and disseminated effects to be in opposite directions, but it is technically possible. Because the overall effect depends on spurious features of the study design, including the size of the networks and the number of index members, it will not be generalizable from one study to the next or to any scaled-up population, unless these features remain constant.

When there is no disseminated effect, the overall will always be less than or equal to the composite. We also show that the overall effect will equal the composite only when the individual effect is null, a rather unlikely occurrence in our motivational setting. In Web Appendix 1, we also show properties for the relationship between the overall and composite risk ratio.

Web Appendix 2 illustrates these relationships through some numerical studies motivated by HPTN 037. If the individual and disseminated effects are in the same direction, the magnitude of the overall effect decreases as the network size increases. In the extreme, when there is no disseminated effect, the overall effect will approach the null as the network size increases, while the composite effect remains constant.

Estimation and Inference for Individual and Disseminated Effects

In network-randomized trials, the overall effect estimate has an immediate causal interpretation. In contrast, index status is not randomized. The indexes are recruited and then the remaining network members are recruited by the index. Hence, the individual, disseminated, and composite effects only have a causal interpretation when the estimator is fully adjusted for confounding.

Generalized estimating equations (GEEs) (17) with a log link and working binomial variance can be used to estimate relative risks or rates, and an identity link and working binomial variance can be used to estimate risk or rate differences, and their confidence intervals adjusted for confounding (18, 19). These models also adjust the estimated parameter variances for correlation within networks and, if the data are longitudinal, across visits within a participant. For log and identity links as employed in this paper, the conditional and marginal model parameters of interest are equivalent because the conditional mean is additive for the fixed and random effects; thus, the estimated effects can be interpreted as either participant-level and/or population-level estimates (20).

One way to estimate these parameters is using an *aggregate* model. Assuming that the effects of the covariates $\boldsymbol{Z}_{ki}$ are not modified by index status R_{ki} and the linear model with the identity link fits the data, let

$$E\left[Y_{kij}\middle|R_{ki}, X_k, \boldsymbol{Z}_{ki}\right] = \gamma_0 + \gamma_1 R_{ki} + \gamma_2 X_k + \gamma_3 X_k R_{ki} + \gamma_4 \boldsymbol{Z}_{ki} .$$

It follows that the effect of being an index member in a control network is $E\left[Y_{kij}\middle|R_{ki} = 1, X_k = 0, \boldsymbol{Z}_{ki}\right] - E\left[Y_{kij}|R_{ki} = 0, X_k = 0, \boldsymbol{Z}_{ki}\right] = \gamma_1$. In a network randomized trial, there could be residual confounding even after adjusting for covariates $\boldsymbol{Z}_{ki}$, so subtracting off these terms accounts for possible unmeasured confounding due to self-selection of index status when estimating individual and composite effects (21, 22). Thus, the individual rate difference (RD) can be estimated by

$$\widehat{RD}_a^I = \hat{E}\left[Y_{kij}\middle|R_{ki} = 1, X_k = 1, \boldsymbol{Z}_{ki}\right] - \hat{E}\left[Y_{kij}\middle|R_{ki} = 0, X_k = 1, \boldsymbol{Z}_{ki}\right]$$
$$-\left\{\hat{E}\left[Y_{kij}\middle|R_{ki} = 1, X_k = 0, \boldsymbol{Z}_{ki}\right] - \hat{E}\left[Y_{kij}|R_{ki} = 0, X_k = 0, \boldsymbol{Z}_{ki}\right]\right\}$$
$$= \hat{\gamma}_0 + \hat{\gamma}_1 + \hat{\gamma}_2 + \hat{\gamma}_3 + \hat{\gamma}_4 \boldsymbol{Z}_{ki} - (\hat{\gamma}_0 + \hat{\gamma}_2 + \hat{\gamma}_4 \boldsymbol{Z}_{ki}) - \hat{\gamma}_1 = \hat{\gamma}_3.$$

When estimating the disseminated effect, only information from non-index network members is included, therefore residual confounding of R_{ki} is not a concern. Adjustment for observed baseline covariates $\boldsymbol{Z}_{ki}$ is needed because randomization in the full study sample does not necessarily guarantee exchangeability of X_k within subgroups of participants. The disseminated RD can be estimated by

$$\widehat{RD}_a^D = \hat{E}\left[Y_{kij}\middle|R_{ki} = 0, X_k = 1, \boldsymbol{Z}_{ki}\right] - \hat{E}\left[Y_{kij}\middle|R_{ki} = 0, X_k = 0, \boldsymbol{Z}_{ki}\right]$$

$$= \hat{\gamma}_0 + \hat{\gamma}_2 + \hat{\gamma}_4\boldsymbol{Z}_{ki} - (\hat{\gamma}_0 + \hat{\gamma}_4\boldsymbol{Z}_{ki}) = \hat{\gamma}_2.$$

Similarly, the composite RD can be estimated by

$$\widehat{RD}_a^{Comp} = \hat{E}\left[Y_{kij}\middle|R_{ki} = 1, X_k = 1, \boldsymbol{Z}_{ki}\right] - \hat{E}[Y_{kij}|R_{ki} = 0, X_k = 0, \boldsymbol{Z}_{ki}]$$

$$-\left\{\hat{E}\left[Y_{kij}\middle|R_{ki} = 1, X_k = 0, \boldsymbol{Z}_{ki}\right] - \hat{E}\left[Y_{kij}|R_{ki} = 0, X_k = 0, \boldsymbol{Z}_{ki}\right]\right\}$$

$$= \hat{\gamma}_0 + \hat{\gamma}_1 + \hat{\gamma}_2 + \hat{\gamma}_3 + \hat{\gamma}_4\boldsymbol{Z}_{ki} - (\hat{\gamma}_0 + \hat{\gamma}_4\boldsymbol{Z}_{ki}) - \hat{\gamma}_1 = \hat{\gamma}_2 + \hat{\gamma}_3.$$

Alternatively, if the effects of covariates $\boldsymbol{Z}_{ki}$ differ by index status R_{ki}, a *stratified* model could be used (Web Appendix 3). The estimators of the risk or rate ratio of the three effects of interest are defined analogously and can be estimated using a GEE with a log link and a working binomial variance. SAS code provided in Web Appendix 4 demonstrates how to obtain these estimates and their corresponding variances. Analyses were performed using SAS Version 9.4 (Cary, NC).

Illustrative Example

The HPTN 037 study (23) was a Phase III randomized controlled HIV preventive intervention trial among people who inject drugs in the United States and Thailand (13). Following a network-randomized design, the index participants were eligible if they reported injecting drugs at least 12 times in the last three months, while the network members had to have injected drugs or had sex with the index member within the last three months. This study assessed the efficacy of a network-oriented peer education intervention to promote HIV risk reduction behaviors among people who inject drugs. Participants were followed for up to 30 months with visits biannually with a median follow-up time of 18 months (Quartile 1 = 6, Quartile 3 = 24) to obtain information on HIV incidence and risk behaviors. The study was underpowered for the primary outcome HIV incidence, so this analysis focused on the occurrence of reported HIV risk behaviors. Two sites participated, Chiang Mai, Thailand and Philadelphia, Pennsylvania, USA. At the time of this study, there was a "war on drugs" in Thailand, which may have reduced trust among people who inject drugs, possibly making the intervention less effective. Therefore, following the recommendation of the study investigators (Carl Latkin, Johns Hopkins Bloomberg School of Public Health, personal communication, 2016), this analysis only included participants at the Philadelphia site. Index participants whose network was randomized to the intervention arm received an educational intervention at baseline and education boosters at six and 12 months. Participants in both the intervention and control arms received HIV counseling and testing at each study visit. The primary analysis for this trial reported the overall effect estimated by a two-level GEE that accounted for correlations between participants within a network and between visits within participants (13).

Shared cotton, an indicator for sharing needle/syringe "works", was selected as an outcome because it nicely exemplified our methods. A more comprehensive clinical outcome, *any injection-related risk* behavior, included the following: sharing injection equipment (needles, cookers, cotton, and rinse water), front and back loading (i.e., injecting drugs from one syringe to another), injected with people not well known or in shooting gallery, and not properly disinfecting injection equipment. Following the original analysis of this study, these outcomes were assessed among participants who reported injection drug use in the last six months at baseline. Statistical tests comparing prevalence of risk behaviors at baseline between network and index members were performed using a GEE model that accounted for within-network correlation.

First, the cumulative incidence of ever reporting the outcome by 30 months of follow-up was analyzed using GEEs to account for correlations within networks, using the robust sandwich estimator with a working exchangeable correlation matrix (24, 25). Next, the longitudinal data were used to assess the effects of the intervention on the inter-visit incidence rates of sharing "works" and of any injection-related risk behavior using a multilevel GEE model. For estimation of the individual and disseminated effects, these models were adjusted for baseline covariates that were known or suspected risk factors for the outcome. For a few outcomes, the models did not converge and log-Poisson models, which provide consistent but not fully efficient estimates of the relative risk, were used (26, 27). All statistical tests performed were two sided.

RESULTS

At the Philadelphia site of the HPTN 037 trial, there were 696 participants and 560 participants had at least one follow-up visit with a total of 1,598 person visits. At this site, 336 (48%) participants were in intervention networks (Table 2). At baseline, participants in intervention networks had a comparable prevalence of reported injection drug risk behavior (84%) compared to those in control networks (86%); however, index participants reported more risk behaviors at baseline (89%) as compared to network members (84%, $P = 0.07$).

Table 3 presents the cumulative incidence of the risk behavior *sharing "works"* and *any report* of injection-related risk behavior. Table 4 presents the results for the effects of the intervention on each outcome. For both outcomes, there was no evidence that the stratified model fit better than the aggregate model on either scale based on an informal comparison of the log likelihoods. Tables 5 and 6 display the results of six-month inter-visit incidence rate differences and ratios for *sharing "works"* and *any report* of injection-related risk behavior from the longitudinal data. Based on an informal comparison of the log likelihoods, the stratified model was a better fit than the aggregate model on the ratio scale for both outcomes and on the difference scale for the any report outcome only.

There was a significant 40% overall reduction on the ratio scale (95% confidence interval (CI): 8%, 61%) for the risk of *ever* sharing "works" by 30 months (Table 4). In contrast, there was a substantially larger, significant 61% reduction on the ratio scale in the adjusted composite risk of *ever* sharing "works" due to the intervention (95% CI: 22%, 80%). The individual effect was nearly twice as large as the disseminated effect, Risk Ratio (RR) = 0.52 vs. RR = 0.76, respectively.

In the longitudinal data, the overall intervention effect was significantly protective with a 44% rate reduction on the ratio scale (95% CI: 11%, 65%) for sharing "works" (Table 5). Based upon the adjusted stratified models, there was a significant protective effect observed among network members with a 41% rate reduction on the ratio scale (95% CI = 3%, 64%), and a somewhat greater 59% individual rate reduction (95% CI = -6%, 84%). A significant 76% composite adjusted rate reduction on the ratio scale was observed (95% CI = 45%, 89%).

For the risk of *ever* any report of risk behavior by 30 months, there was a near null individual effect, while the disseminated effect showed some suggestion of protection (Table 4). In the adjusted aggregate models, there was a non-statistically significant 17% reduction on the ratio scale in the composite risk of any report due to the intervention (95% CI = -10%, 38%) and a comparable overall risk reduction. The longitudinal analysis of the any injection-related risk behavior outcome demonstrated a statistically significant protective overall effect with a 28% rate reduction on the ratio scale (95% CI: 10%, 43%) (Table 6). Based on the adjusted stratified models, the intervention provided a 29% rate reduction on the ratio scale among network members (95% CI = 7%, 46%), but the individual effect did not achieve statistical significance (RR = 0.92, 95% CI = 0.60, 1.40). A significant 35% adjusted composite rate reduction on the ratio scale for any behavior was observed (95% CI = 10%, 53%). As a sensitivity analysis, we used a compound symmetric correlation matrix within a network between subjects and a first order autoregressive correlation matrix within subject over time and the results were comparable to those reported in Tables 5 and 6.

DISCUSSION

We developed estimators for individual and disseminated effects in network-randomized trials. Because networks were randomized to the intervention, the overall effect estimate has a causal interpretation. However, the overall effect is influenced by ancillary factors, such as the size of the networks, and will typically underestimate the composite effect (See Web Appendices 1 and 2). When there is no unmeasured confounding and the model is correctly specified, individual and disseminated effect estimates also have causal interpretations and provide a more in-depth understanding of the intervention's impact.

In the HPTN trial, the overall effect of the intervention was statistically significant with an estimated 28% risk reduction of any injection-related risk behavior; however, although there was evidence for a significant 29% disseminated risk reduction, the individual effect did not achieve statistical significance for that same outcome. The original investigators reported only the overall effects (13), which we found were slightly to moderately attenuated compared to the composite effects that reveal the full power of the intervention. Somewhat surprisingly, the disseminated effect was stronger than the individual effect for the report of any injection-related risk behavior, suggesting that this intervention has substantial resonance within the network beyond the effect of directly receiving the intervention. Without consideration of dissemination, efforts to understand the full array of mechanisms by which the intervention achieved its goal would be likely overlooked.

The assumption of no unmeasured covariates associated with the treatment and outcome (or with the index status and outcome) cannot be empirically verified. For example, in HPTN 037, an individual's unmeasured communication skills may affect whether or not they come forward to be an index and this may lead to unmeasured differences between index and non-index members. Future work could involve extensions to address unmeasured confounding when evaluating disseminated effects. In addition, the methods for incidence rate measures assume that there is no bias due to dependent loss-to-follow-up, and in the longitudinal analysis, the missing visit process is ignorable (i.e., missing a visit is independent of the outcome conditional on intervention status and observed baseline covariates). If this assumption is questionable, censoring weights could be employed in the analysis (28, 29). These methods assume no dissemination between networks, although, it is possible in some settings that some network members will be in the risk networks of more than one index member. In HPTN 037, indexes may interact with participants outside their observed risk network because they may frequent the same neighborhoods and venues. These methods could thus be extended to accommodate dissemination between as well as within networks. In the HPTN trial, effect modification was observed for sex and participation in a drug treatment program. Future work could entail estimation of these within-strata effects and an application of such methods as g-estimation to ascertain population-level effects (30).

These methods could also be extended to correct for bias due to misclassification or measurement error in the self-reported outcome or covariates (31). We assume that the reported effect estimates are not subject to social desirability effects, which could vary by intervention arm over time and dissemination may reinforce this. Furthermore, the indexes may have underreported risk connections or study investigators may have missed some networks entirely. More accurate ways to elicit and recruit network member nominations and contact information could be developed and methods to infer unobserved or misclassified risk and social connections could be improved.

A network-based program implementation can be offered at a reduced cost, because only a subset of participants needs to receive the intervention. The example highlights the need for methods to adequately power trials to assess individual and disseminated effects. Future work could include evaluating the disseminated effects of treatment as prevention and similar interventions in HIV trials, including extensions for networks with more than one index participant (32-36). Extension of these methods to estimate both individual and disseminated effects of the components of multifaceted interventions is needed for future complex HIV/AIDS implementation science research, particularly that engaging drug-using or sexual risk networks.

Acknowledgements

Author affiliations: Department of Pharmacy Practice, College of Pharmacy, University of Rhode Island, Kingston, Rhode Island (Ashley L. Buchanan); Vanderbilt Institute for Global Health and Department of Pediatrics, Vanderbilt University School of Medicine, New Haven, Connecticut (Sten H. Vermund); Institute for Infectious Disease Research, National Development and Research Institutes, Inc., New York, New York (Samuel R. Friedman); Departments of Epidemiology, Biostatistics, Nutrition, and Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA (Donna Spiegelman).

We thank Drs. Deborah Donnell and Carl Latkin for access to the HIV Prevention Trials Network (HPTN) 037 data and for valuable suggestions. We thank Dr. James Robins for his expert advice. Data from the HPTN were obtained with support from the National Institute on Drug Abuse (NIDA), National Institute of Allergies and Infectious Disease (NIAID), and National Institute of Mental Health (NIMH) under National Institutes of Health (NIH) grants UM1AI068619 (HPTN Leadership and Operations Center), UM1AI068617 (HPTN Statistical and Data Management Center), and UM1AI068613 (HPTN Laboratory Center). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors have no conflicts of interest to declare.

References

1. VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York, NY: Oxford University Press; 2015.
2. Rubin DB. Bayesian inference for causal effects: The role of randomization. *Ann Stat*. 1978;7:34-58.
3. Hudgens MG, Halloran ME. Toward Causal Inference With Interference. *J Am Stat Assoc*. 2008;103(482):832-842.
4. Liu L, Hudgens MG. Large sample randomization inference of causal effects in the presence of interference. *J Am Stat Assoc*. 2014;109(505):288-301.
5. Halloran ME, Struchiner CJ. Study designs for dependent happenings. *Epidemiology*. 1991;2(5):331-338.
6. Struchiner CJ, Halloran ME, Robins JM, Spielman A. The behaviour of common measures of association used to assess a vaccination programme under complex disease transmission patterns—a computer simulation study of malaria vaccines. *Int J Epidemiol*. 1990;19(1):187-196.
7. Tchetgen Tchetgen EJ, VanderWeele TJ. On causal inference in the presence of interference. *Stat Methods Med Res*. 2012;21(1):55-75.
8. Perez-Heydrich C, Hudgens MG, Halloran ME, Clemens JD, Ali M, Emch ME. Assessing effects of cholera vaccination in the presence of interference. *Biometrics*. 2014;70(3):731-744.
9. Lundin M, Karlsson M. Estimation of causal effects in observational studies with interference between units. *Stat Methods Appt*. 2014;23(3):417-433.
10. Friedman SR, Bolyard M, Khan M, Maslow C, Sandoval M, Mateu-Gelabert P, Krauss B, Aral SO. Group sex events and HIV/STI risk in an urban network. *J Acquir Immune Defic Syndr*. 2008;49(4):440-446.
11. Khan MR, Epperson MW, Mateu-Gelabert P, Bolyard M, Sandoval M, Friedman SR. Incarceration, sex with an STI-or HIV-infected partner, and infection with an STI or HIV in Bushwick, Brooklyn, NY: A social network perspective. *Am J Public Health*. 2011;101(6):1110-1117.
12. Friedman S, Bolyard M, Sandoval M, Mateu-Gelabert P, Maslow C, Zenilman J. Relative prevalence of different sexually transmitted infections in HIV-discordant sexual partnerships: Data from a risk network study in a high-risk New York neighbourhood. *Sex Transm Infect*. 2008;84(1):17-18.
13. Latkin CA, Donnell D, Metzger D, Sherman S, Aramrattna A, Davis-Vogel A, Quan VM, Gandham S, Vongchak T, Perdue T. 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(4):740-748.
14. Williams C, Eisenberg M, Becher J, Davis-Vogel A, Fiore D, Metzger D. Racial disparities in HIV prevalence and risk behaviors among injection drug users and members of their risk networks. *J Acquir Immune Defic Syndr*. 2013;63:S90-S94.
15. Latkin C, Donnell D, Liu TY, Davey-Rothwell M, Celentano D, Metzger D. The dynamic relationship between social norms and behaviors: The results of an HIV prevention network intervention for injection drug users. *Addiction*. 2013;108(5):934-943.

16. Ogburn EL, VanderWeele TJ. Causal diagrams for interference. *Stat Sci*. 2014;29(4):559-578.
17. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis*: Hoboken, New Jersy: John Wiley and Sons; 2012.
18. Wacholder S. Binomial regression in GLIM: estimating risk ratios and risk differences. *Am J Epidemiol*. 1986;123:174-184.
19. Deddens J, Petersen MR, Endahl L. Prevalence proportion ratios: estimation and hypothesis testing. *Int J Epidemiol*. 1998;27(1):91-95.
20. Ritz J, Spiegelman D. Equivalence of conditional and marginal regression models for clustered and longitudinal data. *Stat Methods in Med Res*. 2004;13(4):309-323.
21. Meyer BD. Natural and quasi-experiments in economics. *Journal of Business and Economic Statistics*. 1995;13(2):151-161.
22. Lipsitch M, Tchetgen ET, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21(3):383.
23. Sista ND, Karim QA, Hinson K, Donnell D, Eshleman SH, Vermund SH. Experience in international clinical research: the HIV Prevention Trials Network. *Clin Investig*. 2011;1(12):1609-1618.
24. Huber P, editor. The behavior of maximum likelihood estimates under nonstandard conditions. Fifth Berkeley Symposium on Mathematical Statistics and Probability; 1967; Berkely, CA: University of California Press.
25. Liang KY, Zeger SL. Longitudinal analysis using generalized linear models. *Biometrika*. 1986;73(1):13-22.
26. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159(7):702-706.
27. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am J Epidemiol*. 2005;162(3):199-200.
28. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted log rank tests. *Biometrics*. 2000:779-788.
29. Rotnitzky A, Robins JM. Semiparametric regression estimation in the presence of dependent censoring. *Biometrika*. 1995;82(4):805-820.
30. Robins JM, Blevins D, Ritter G, Wulfsohn M. G-estimation of the effect of prophylaxis therapy for Pneumocystis carinii pneumonia on the survival of AIDS patients. *Epidemiology*. 1992;3(4):319-336.
31. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. *Measurement Error in Nonlinear Models*. Second ed. London, England: Chapman and Hall; 2006.
32. Kurth AE, Celum C, Baeten JM, Vermund SH, Wasserheit JN. Combination HIV prevention: significance, challenges, and opportunities. *Curr HIV/AIDS Rep*. 2011;8(1):62-72.
33. Tanser F, Barnighausen T, Grapsa E, Zaidi J, Newell ML. High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. *Science*. 2013;339(6122):966-971.
34. Vermund SH, Fidler SJ, Ayles H, Beyers N, Hayes RJ, Team HS. Can combination prevention strategies reduce HIV transmission in generalized epidemic settings

in Africa? The HPTN 071 (PopART) study plan in South Africa and Zambia. *J Acquir Immune Defic Syndr*. 2013;63(2):S221.
35. Hayes R, Ayles H, Beyers N, Sabapathy K, Floyd S, Shanaube K, Bock P, Griffith S, Moore A, Watson-Jones D. HPTN 071 (PopART): Rationale and design of a cluster-randomised trial of the population impact of an HIV combination prevention intervention including universal testing and treatment–a study protocol for a cluster randomised trial. *Trials*. 2014;15(1):57.
36. Chamie G, Kwarisiima D, Clark TD, Kabami J, Jain V, Geng E, Petersen ML, Thirumurthy H, Kamya MR, Havlir DV. Leveraging rapid community-based HIV testing campaigns for non-communicable diseases in rural Uganda. *PloS one.* 2012;7(8):e43400.

Figure 1. Schematic diagram of the subsets of data used for each estimator (individual, disseminated, composite and overall) based on a format provided in (5)

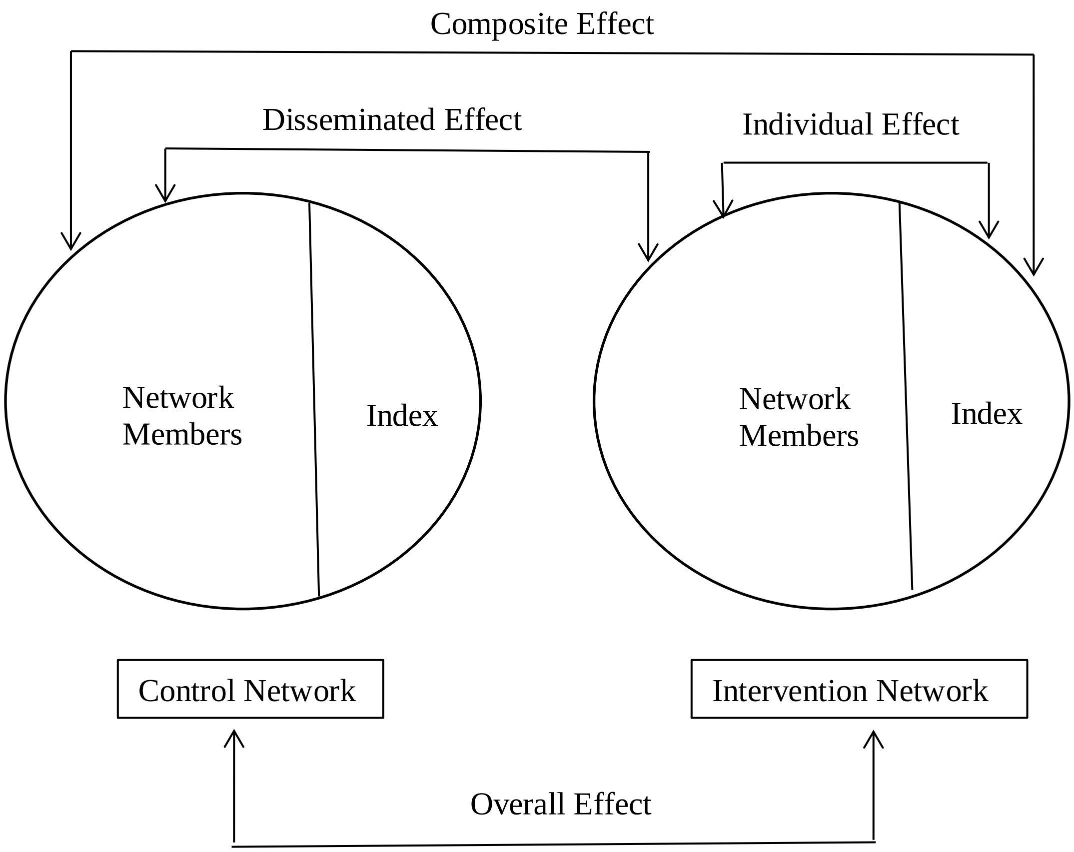

Table 1. Related Terminology for Estimation of Individual and Disseminated Effects

Recommended Term	Alternative Terms	Definition	Parameter of Interest	Network-Randomized Design Estimators	
				Aggregate[a]	Stratified[b]
Individual	Direct	Effect on those directly receiving intervention beyond being in an intervention network	$E\left[Y_{kij}(1,1)\right]^{c} - E\left[Y_{kij}(0,1)\right]^{d}$	$\hat{\gamma}_3$	$\hat{\alpha}_1 - \hat{\beta}_1$
Disseminated	Indirect, Social Diffusion, Diffusion of Innovation, Contamination, Spillover	Effect on those who received intervention indirectly through the index participant	$E[Y_{kij}(0,1)] - E[Y_{kij}(0,0)]^{e}$	$\hat{\gamma}_2$	$\hat{\beta}_1$
Composite	Total	Combined individual and disseminated effects; Effect among indexes in intervention networks contrasted with effect among network members in control networks	$E[Y_{kij}(1,1)] - E[Y_{kij}(0,0)]$	$\hat{\gamma}_2 + \hat{\gamma}_3$	$\hat{\alpha}_1$
Overall[f]	Crude	Effect among members of intervention networks contrasted with effect among members of control networks	$E\left[Y_{kij}(\cdot,1)\right]^{g} - E\left[Y_{kij}(\cdot,0)\right]^{h}$	$\hat{\beta}_1^{*}$	

[a] For a network-randomized design, rate difference parameters of the individual, disseminated, and composite effects, respectively, are estimated from an aggregate GEE model with an identity link and binomial variance:

$$E\left[Y_{kij}\middle|R_{ki}, X_k, \boldsymbol{Z}_{ki}\right] = \gamma_0 + \gamma_1 R_{ki} + \gamma_2 X_k + \gamma_3 X_k R_{ki} + \gamma_4 \boldsymbol{Z}_{ki}$$

[b] For a network-randomized design, rate difference parameters of the individual, disseminated, and composite effects, respectively, are estimated from a stratified GEE model with an identity link and binomial variance:

$$E\left[Y_{kij}\middle|R_{ki}, X_k, \boldsymbol{Z}_{ki}\right] = I(R_{ki} = 0) \times (\beta_0 + \beta_1 X_k + \beta_2 \boldsymbol{Z}_{ki}) + I(R_{ki} = 1) \times (\alpha_0 + \alpha_1 X_k + \alpha_2 \boldsymbol{Z}_{ki})$$

[c] $Y_{kij}(1,1)$ is the potential outcome of participant i at visit j in network k, if possibly contrary to fact, this participant was an index member in a network randomized to the intervention. $E[X]$ is the expectation of the random variable X.

[d] $Y_{kij}(0,1)$ is the potential outcome of participant i at visit j in network k, if possibly contrary to fact, this participant was a network member in a network randomized to the intervention.
[e] $Y_{kij}(0,0)$ is the potential outcome of participant i at visit j in network k, if possibly contrary to fact, this participant was a network member in a network randomized to the control.
[f] For network-randomized design, parameter of the overall effect is estimated from a GEE model with an identity link and binomial variance:

$$E[Y_{kij}|X_k] = \beta_0^* + \beta_1^* X_k$$

[g] $Y_{kij}(\cdot\,,1)$ is the potential outcome of participant i at visit j in network k, if possibly contrary to fact, this participant was in a network randomized to the intervention.
[h] $Y_{kij}(\cdot\,,0)$ is the potential outcome of participant i at visit j in network k, if possibly contrary to fact, this participant was in a network randomized to the control.

Table 2 Baseline characteristics of the HPTN 037 study population, Philadelphia site, 2002-2006, by treatment group (n = 696)

	Network Randomization[a]	
Characteristics	Intervention % of (n = 336)	Control % of (n = 360)
Network Member Role		
Index	33	33
Network member	67	67
Network size[b]	3 (1.27)	3 (1.04)
Male	68	69
Age (years)[b]	41 (10)	40 (10)
Drug treatment program[b]	22	32
Housing[c]		
Spent night on street, car, park or abandoned building	25	24
Spent time in jail	19	15
Sexual risk[d]		
More than one sex partner	42	40
Unprotected sex in the last week	50	50
Unprotected sex with non-primary Partner	16	19
Injection drug behaviors[e]		
Injection drug use in the last 6 month	93	94
Injection drug use in the last month	99	97
Heroin	94	96
Heroin w/ Cocaine	42	35
Heroin w/ Amphetamine	1	1
Cocaine	37	37
Amphetamine	2	2
Number of days injected in the last month		
0-5	9	9
6-14	8	9
15-29	24	16
Everyday	59	65
Injection risk behaviors in the last month[f]		
Shared rinse water	44	51
Shared cooker	57	62
Shared cotton	41	46
Used front or back loaded syringe	21	23
Injected with an unclean syringe	13	15
Passed a syringe to someone else	48	53
Used a syringe after someone else	37	39
Injected with someone you don't know	19	22

well		
Injected in a shooting gallery	61	56
Did not clean syringe after use	14	13

[a] There were 112 intervention networks and 120 control networks. Values of polytomous variables may not sum to 100% due to rounding.
[b] Values are expressed as mean (standard deviation).
[c] In past 6 months.
[d] In last month.
[e] Injection drug behaviors reported only for participants reporting injection drug use in the past 6 months.
[f] Injection risk behaviors reported only for participants reporting injection drug use in the last month.

Table 3. Risk of *ever* reporting injection-related risk behavior by 30 months after baseline with 95% confidence intervals (CI) at the HPTN 037 Philadelphia site, 2002-2006 (n = 651)

Network Randomization	Index Participants			Network Participants		
	Total Person	Shared "Works" Reported	30-month Cumulative Incidence, %	Total Person	Shared "Works" Reported	30-month Cumulative Incidence, %
Any shared works						
Treatment	112	10	9	202	28	14
Control	120	24	20	217	42	19
Combined Control	337	66	20			
Any risk behavior						
Treatment	112	48	43	202	75	37
Control	120	58	48	217	97	45
Combined Control	337	155	46			

Table 4. Effects of the intervention on the risk of *ever* reporting injection-related risk behavior by 30 months after baseline with 95% confidence intervals (CI) at the HPTN 037 Philadelphia site, 2002-2006 (n = 651)

	Stratified Models								Aggregate Models							
	Unadjusted		Adjusted[a]		Unadjusted		Adjusted[a]		Unadjusted		Adjusted[a]		Unadjusted		Adjusted[a]	
Effect	RD	95% CI	RD	95% CI	RR	95% CI	RR	95% CI	RD	95% CI	RD	95% CI	RR	95% CI	RR	95% CI
Any shared works																
Individual	−5	−16, 7	−7	−20, 5	0.65	0.28, 1.51	0.47	0.18, 1.21	−5	−16, 6	−8	−19, 3	0.65	0.28, 1.47	0.52	0.23, 1.15
Disseminated	−6	−14, 1	−5	−13, 2	0.69	0.42, 1.13	0.75	0.45, 1.25	−6	−14, 1	−5	−13, 3	0.69	0.42, 1.14	0.76	0.46, 1.25
Composite	−11	−20, −2	−13	−22, −3	0.45	0.22, 0.89	0.35	0.16, 0.78	−11	−20, −2	−13	−22, −4	0.45	0.22, 0.89	0.39	0.20, 0.78
Overall	−8	−14, −2	−8	−14, −2	0.60	0.39, 0.92	0.61	0.39, 0.94	−8	−14, −2	−8	−14, −2	0.60	0.39, 0.92	0.61	0.39, 0.94
Any risk behavior[b]																
Individual	2	−14, 18	0	−18, 17	1.07	0.73, 1.56	0.99	0.68, 1.45	2	−13, 18	0	−16, 15	1.07	0.76, 1.53	0.97	0.68, 1.39
Disseminated	−8	−18, 2	−5	−16, 5	0.83	0.64, 1.07	0.88	0.69, 1.13	−8	−18, 2	−7	−17, 4	0.83	0.64, 1.06	0.85	0.66, 1.10
Composite	−5	−18, 7	−6	−20, 8	0.89	0.67, 1.18	0.88	0.66, 1.16	−5	−18, 7	−7	−20, 6	0.89	0.67, 1.18	0.83	0.62, 1.10
Overall	−7	−15, 1	−7	−16, 2	0.85	0.69, 1.04	0.85	0.69, 1.04	−7	−15, 1	−7	−16, 2	0.85	0.69, 1.04	0.85	0.69, 1.04

CI = Confidence Interval; RD = Risk Difference per 100 persons; RR = Risk Ratio.

[a] Adjusted for sex (male vs. female), age (years), marital status (single vs. not single), education (at least high school vs. not), and employment (unemployed vs. employed), and time-varying covariates set to their baseline value: crack use (yes vs. no), cocaine use (yes vs. no), benzodiazepines (yes vs. no), heroin (yes vs. no), drug treatment program (yes vs. no), spent the night on the street (yes vs. no), spent time in jail (yes vs. no), alcohol use (got drunk vs. not), injected heroin (yes vs. no), heroin and cocaine (yes vs. no), injected cocaine (yes vs. no), and number of days injected in the last month (0-5 days, 6-14 days, 15-29 days vs. everyday).

[b] On the ratio scale, the model excluded number of days injected variable because of model convergence issues. Models for the overall effect also excluded injected cocaine. Stratified model for individual, disseminated, and composite effects also excluded spent time in jail, heroin use, and injected cocaine.

Table 5. Six-month inter-visit incidence rate ratios and rate differences for the effect of the HPTN 037 randomized intervention on the rate of *sharing "works"* risk behavior during follow-up with 95% confidence intervals (CI) among participants with at least one follow-up visit at the Philadelphia site, 2002-2006

	Unadjusted				Baseline Adjusted[a,b]			
	RR	95% CI	RD	95% CI	RR	95% CI	RD	95% CI
				Stratified Models				
Overall[c]	0.56	0.35, 0.89	-0.05	-0.10, -0.01	0.52	0.32, 0.83	-0.06	-0.11, -0.01
Individual[d]	0.58	0.23, 1.48	-0.05	-0.14, 0.04	0.41	0.16, 1.06	-0.08	-0.18, 0.01
Disseminated[e]	0.66	0.38, 1.12	-0.04	-0.09, 0.02	0.59	0.36, 0.97	-0.04	-0.09, 0.02
Composite[f]	0.38	0.18, 0.81	-0.09	-0.16, -0.02	0.24	0.11, 0.55	-0.12	-0.19, -0.05
				Aggregate Models				
Overall	0.56	0.35, 0.89	-0.05	-0.10, -0.01	0.52	0.32, 0.83	-0.06	-0.11, -0.01
Individual	0.56	0.25, 1.26	-0.06	-0.13, 0.02	0.55	0.25, 1.19	-0.08	-0.15, 0.00
Disseminated	0.69	0.04, 1.17	-0.03	-0.09, 0.02	0.64	0.38, 1.08	-0.03	-0.09, 0.02
Composite	0.39	0.18, 0.81	-0.09	-0.16, -0.02	0.35	0.17, 0.73	-0.11	-0.18, -0.04

CI = Confidence Interval; RD = Risk Difference per 100 person-visits; RR = Risk Ratio.
[a] Adjusted for sex (male vs. female), age (years), marital status (single vs. not single), education (at least high school vs. not), and employment (unemployed vs. employed), and time-varying covariates set to their baseline value: crack use (yes vs. no), cocaine use (yes vs. no), benzodiazepines (yes vs. no), smoked heroin (yes vs. no), drug treatment program (yes vs. no), spent the night on the street (yes vs. no), spent time in jail (yes vs. no), alcohol use (got drunk vs. not), injected heroin (yes vs. no), heroin and cocaine (yes vs. no), injected cocaine (yes vs. no), and number of days injected in the last month (0-5 days, 6-14 days, 15-29 days vs. everyday).
[b] One participant missing information on spent the night on the street and spent time in jail at baseline.
[c] There were 174 events, 1,598 person-visits and 560 people included.
[d] There were 58 events, 782 person-visits and 270 people included.

[e] There were 158 events, 1,319 person-visits and 463 people included.
[f] There were 132 events, 1,095 person-visits and 387 people included.

Table 6. Six-month inter-visit incidence rate ratios and rate differences for the effect of the HPTN 037 randomized intervention on the rate of *any injection-related risk* behavior during follow-up with 95% confidence intervals (CI) among participants with at least one follow-up visit at the Philadelphia site, 2002-2006

	Unadjusted				Baseline Adjusted[a,b]			
	RR	95% CI	RD	95% CI	RR	95% CI	RD	95% CI
Stratified Models								
Overall[c]	0.72	0.57, 0.90	-0.09	-0.16, -0.02	0.69	0.56, 0.86	-0.09	-0.16, -0.03
Individual[d]	0.97	0.64, 1.48	-0.02	-0.15, 0.12	0.92	0.60, 1.40	-0.06	-0.19, 0.07
Disseminated[e]	0.73	0.55, 0.97	-0.09	-0.17, 0.00	0.71	0.54, 0.93	-0.08	-0.16, -0.01
Composite[f]	0.71	0.52, 0.97	-0.10	-0.21, 0.00	0.65	0.47, 0.90	-0.14	-0.24, -0.03
Aggregate Models								
Overall	0.72	0.57, 0.90	-0.09	-0.16, -0.02	0.69	0.56, 0.86	-0.09	-0.16, -0.03
Individual	0.96	0.67, 1.39	-0.02	-0.14, 0.10	0.91	0.62, 1.33	-0.05	-0.17, 0.07
Disseminated	0.75	0.57, 0.98	-0.08	-0.16, 0.00	0.72	0.55, 0.94	-0.08	-0.16, 0.01
Composite	0.72	0.53, 0.98	-0.10	-0.21, 0.01	0.65	0.47, 0.90	-0.13	-0.23, -0.02

CI = Confidence Interval; RD = Risk Difference per 100 person-visits; RR = Risk Ratio.

[a] Adjusted for sex (male vs. female), age (years), marital status (single vs. not single), education (at least high school vs. not), and employment (unemployed vs. employed), and time-varying covariates set to their baseline value: crack use (yes vs. no), cocaine use (yes vs. no), benzodiazepines (yes vs. no), smoked heroin (yes vs. no), drug treatment program (yes vs. no), spent the night on the street (yes vs. no), spent time in jail (yes vs. no), alcohol use (got drunk vs. not), injected heroin (yes vs. no), heroin and cocaine (yes vs. no), injected cocaine (yes vs. no), and number of days injected in the last month (0-5 days, 6-14 days, 15-29 days vs. everyday).

[b] One participant missing information on spent the night on the street and spent time in jail at baseline.

[c] There were 509 events, 1,598 person-visits and 560 people included.

[d] There were 204 events, 782 person-visits and 270 people included.

[e] There were 433 events, 1,319 person-visits and 463 people included.
[f] There were 381 events, 1,095 person-visits and 387 people included.