Harvard University

Harvard University Biostatistics Working Paper Series

Year 2013 *Paper* 161

Sample Size Considerations in the Design of Cluster Randomized Trials of Combination HIV Prevention

Rui Wang* Ravi Goyal† Quanhong Lei‡

M. Essex** Victor DeGruttola††

*Harvard School of Public Health and Brigham and Women's Hospital, rwang@hsph.harvard.edu

†Harvard School of Public Health, rgoyal@shph.harvard.edu

‡Harvard School of Public Health, lei@hsph.harvard.edu

**Harvard School of Public Health, messex@hsph.harvard.edu

††Harvard School of Public Health, degrut@hsph.harvard.edu

Sample size considerations in the design of cluster randomized trials of combination HIV prevention

Rui Wang[a,b], Ravi Goyal[a], Quanhong Lei[a], M. Essex[c], Victor De Gruttola[a]

[a]Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA;

[b]Division of Sleep Medicine, Brigham and Women's Hospital, Boston, MA, USA;

[c]Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA, USA;

Running Title: Design of cluster randomized trials

Author for correspondence: Rui Wang, Division of Sleep Medicine, Brigham and Women's Hospital, 221 Longwood Ave., Room 255, Boston, MA 02115. Email: rwang@hsph.harvard.edu.

Acknowledgment: This research was supported by R01 AI24643 and R01 AI51164 from the National Institutes of Health, and U01 GH000447 from the Centers for Disease Control and Prevention.

Background. Cluster randomized trials have been utilized to evaluate the effectiveness of HIV prevention strategies when the study endpoint is population-level incidence (Hughes and Kulich 2006). Design of such studies must take into account possible correlation of outcomes within randomized units.

Purpose. This paper discusses the power and sample size considerations for cluster randomized trials of combination HIV prevention, using a proposed HIV prevention study in Botswana as an illustration.

Methods. We introduce a new agent-based simulation model to evaluate the impact of combination prevention strategy at a community level and investigate how correlation structure within a community affects the coefficient of variation–an essential parameter in designing a cluster randomized trial.

Results. We constructed collections of sexual networks and then propagated HIV on these networks to simulate the disease epidemic. Increasing level of sexual mixing between intervention and standard-of-care communities reduces the difference in cumulative incidence in the two sets of communities. Fifteen clusters per arm and 500 incidence cohort members per community provides 96% power to detect the projected difference in cumulative HIV incidence between SOC and intervention communities (3.82% and 2.24%) at the end of the third study year, using the model-based projected value of coefficient of variation 0.24. Although available formulas for calculating sample size can be derived from random effects models in which cluster-level random effects are assumed to be independent across clusters, as are individual outcomes within clusters (e.g., an exchangeable correlation structure), we show that deviations from an exchangeable correlation assumption do not generally affect the validity of such formulas.

Limitations. We construct our sexual network based on data from Likoma island, Malawi and base disease progression on longitudinal estimates from an incidence cohort in Botswana and in Durban as well as a household survey in Mochudi, Botswana. It would be desirable to have alternative network data and more precise estimates based on larger sample sizes for

disease progression to assess the robustness of our model results.

Conclusions. Epidemic modeling plays a critical role in planning and evaluating treatment for prevention. Simulation studies allow us to take into consideration available information on sexual network characteristics, such as mixing within and between communities, as well as coverage levels for different prevention modalities in the combination prevention package under study.

Keywords: cluster randomized trials; network models; design effect; HIV prevention

Background

Many HIV transmission prevention approaches have been demonstrated to be effective when studied separately in individual-level clinical trials, and efforts are now underway to investigate whether combining these modalities can achieve community-level control of HIV infection (Boily et al., 2012). The prevention modes thus far demonstrated to be efficacious include antiretroviral treatment as prevention (TasP), male circumcision (MC), pre-exposure prophylaxis (PrEP) (in some populations) and preventing mother-to-child transmission (PMTCT). The underlying rationale for TasP is the increase in risk of HIV transmission with increases in viral load. Quinn et al. (2000) estimated that each ten-fold increment in viral load is associated with a 2.45-fold increase in the risk of HIV transmission per sexual contact. Gray et al. (2001) reported that the transmission rate per couple increased from 1/43 to 11/45 comparing subjects with viral load < 1700 copies/ml to those with viral load between $1,700$ copies/ml to $12,400$ copies/ml.

Incidence of infection depends on factors operating both at the individual subject level, such as risk behavior, and at the community level, such as sexual network characteristics, i.e., the risk behavior of an individual's partner affects the individual's risk. Nonetheless, prevention strategies target individual-subject characteristics, like viral load of infected partners, to reduce transmission efficiency per sexual act have the potential for significant reductions in incidence (Granich et al., 2009). A meta-analysis based on 5021 heterosexual couples from 11 cohorts reported a 92% reduction in heterosexual transmission rates comparing patients on ART to those not on ART, and no transmission events were observed in discordant heterosexual couples if the HIV-infected partner was treated with ART and had a viral load below 400 copies/ml (Attia et al., 2009). Recently, a randomized clinical trial conducted by HIV Prevention Trials Network (HPTN Study 052) reported a 96% (95% CI) reduction in HIV transmission in HIV-1 sero-discordant couples comparing immediate versus delayed use of antiretroviral treatment by HIV-infected individuals with a CD4 cell count between 350-

550 per cubic millimeter (cells/mm3) at enrollment and therefore did not require treatment according to treatment guidelines (Cohen et al., 2011).

A modified treatment-for-prevention strategy, proposed by Novtisky et al. (2010), expands treatment to individuals with high viral load, but who otherwise do not qualify for treatment under the current guidelines. Previous studies (Novitsky et al. 2010, 2011) showed that about 25% of new HIV-1C infections in southern Africa maintain high viral load levels for at least 1-2 years; furthermore, those with higher viral load levels appeared to have a quicker CD4 cell count decline. Identifying and treating this subset can lead to both individual benefits, achieved by delaying onset of clinical AIDS, and public health benefits resulting from reductions in HIV transmission (Novitsky and Essex 2012). These considerations have led to interest in evaluating a combination prevention package that includes treatment of high viral load carriers in a cluster-randomized trial in Botswana. Botswana is an appropriate setting for this study, because it has one of the highest rates of HIV prevalence, estimated at 24.8% of adults by UNAIDS in 2009. In addition, the country has had an active program for ART treatment based on CD4 <350 cells/ml and/or AIDS-defining illness as well as high uptakes for the PMTCT program. Rates of male circumcision (MC) have been estimated at to be low (12-14%).

Cluster randomized trials are ideally suited to investigate both direct and indirect effects of prevention interventions. Design of such studies must take into account possible correlation of outcomes within randomized units, which induces between-community variation. In this paper, we describe power and sample size considerations for cluster randomized trials of combination HIV prevention, using the proposed Botswana study as an illustration. We introduce a new agent-based simulation model to evaluate the impact of combination prevention strategy at a community level and investigate how correlation structure within a community may affect the coefficient of variation– an essential parameter in designing a

cluster randomized trial.

Methods

Study design overview

The proposed study is intended to assess whether implementation of a combination of prevention interventions can significantly reduce HIV incidence. Villages in Botswana will be randomized into one of the two arms:

A) "standard of care" (SOC) with ART for HIV-infected individuals with CD4<350 cells/ml or AIDS;

B) ART for HIV-infected individuals with CD4<350 cells/ml or AIDS, plus ART for those with high viral load (>10,000 copies/ml), plus combination prevention: strategies of enhanced HTC, PMTCT, enhanced test-linked care (TLC) in relation to ART initiation and follow-up, and male circumcision.

Sampling strategy

The primary estimate of HIV incidence will be obtained from a cohort identified through a random sample of 20% of households in each community. This cohort (designated incidence cohort) will include all consenting eligible HIV-negative household members. Subjects in the incidence cohort will be followed for 3 years and tested annually for HIV status. In establishing the percentage of households to be sampled, two factors were taken into account: 1) adequacy of power, and 2) minimizing home-based testing in the SOC communities. As enhancing HIV testing is one of the interventions under study, home-based testing coverage in SOC communities will attenuate treatment differences between SOC and intervention communities. We chose 20% based on the power it provides and on estimated treatment effects that are described below.

Matching factors

Because the number of randomized units in a cluster randomized trial is generally small compared to that in a randomized trial of independent units, a pair-matched design helps to ensure that the baseline community characteristics are similar for communities randomized into the 2 arms. The communities can be paired based on similarities in population and geographical location. In the Botswana study, the communities cannot be matched based on their baseline incidence because accurate incidence estimates are not available for the proposed study sites.

Sample size determination

Sample size was based on estimates of power to detect intervention effects calculated from the following formula developed for matched cluster randomized trials (Hayes and Moulton 2009):

$$c = 2 + (z_{\alpha/2} + z_{\beta})^2 \frac{\pi_0(1-\pi_0)/m + \pi_1(1-\pi_1)/m + k_m^2(\pi_0^2 + \pi_1^2)}{(\pi_0 - \pi_1)^2},$$

where c is the number of clusters required per treatment arm, π_0 and π_1 are the true proportions of who reach endpoint in the two treatment arms, respectively; m is the number of individuals within each cluster, k_m is the coefficient of variation in true proportions between clusters within matched pairs in the absence of intervention, and $z_{\alpha/2}$ and z_{β} are the usual standard normal distribution values corresponding to upper tail probabilities of $\alpha/2$ and β, respectively. Although our endpoint will be interval-censored times to infection, we develop power calculations for simplified binary outcomes. These should provide a reasonable approximation to power, but may be somewhat conservative as they do not make use of times to event.

To predict the cumulative incidence over the study period in intervention and control communities, we used an agent-based epidemic simulation model, in which the HIV was propagated on collections of generated sexual networks. Parameter values in the model were set based on published literature as well as information from (1) the Mochudi study, a pilot study to

evaluate the uptake of an HIV prevention program for the northeast sector of Mochudi, a village in Botswana with a population of around 45,000 (Botswana Population and Housing Census 2011); (2) the Botswana/Durban cohort, a cohort of acutely or recently infected individuals combined from two southern African cohorts: the HIV pathogenesis Programme Acute Infection Study in Durban, KwaZulu-Natal, South Africa (Wright et al., 2011), and the Tshedimoso Study in Gaborone, Botswana (Novitsky et al., 2009; Novitsky et al., 2008; Novitsky et al., 2011); and (3) the Likoma Island sexual network, a cross-sectional socio-centric survey of sexual partnerships aiming to investigate the population-level structure of sexual networks connecting the young adult population of several villages on Likoma Island, Malawi (Helleringer and Kohler 2007).

Results

Generation of sexual networks.

In our models, sexual networks are represented as random graphs, in which each node represents an individual and each edge, a sexual relationship between the two individuals portrayed by the endpoints. These networks contain 2 types of nodes, male and female. The networks are bipartite; that is, we only consider relationships between opposite genders, reflecting the belief that the principal mode of transmission in Botswana is heterosexual transmission (National AIDS Coordinating Agency, 2010). Furthermore, HIV transmission from homosexual contact cannot be easily documented in Botswana. Each network represents all of the sexual relationships that occur in sets of matched pairs of communities during the course of study. An illustrative network of 2 communities is provided in Figure 1.

[Figure 1 about here.]

In a sexual contact network, the number of edges adjacent to a particular node is called its degree, and the degree distribution can be obtained by the collection of nodal degrees (Wasserman and Faust 1994). We construct degree distributions from the reported number

of sexual partners in four years from Likoma Island using a likelihood approach based on the Yule distribution (Jones and Handcock 2003, Handcock et al., 2003).

Using the methods proposed in Goyal et al. (2011) that permit incorporation of user-specified uncertainty associated with particular network properties, we generated networks that are consistent with both a prescribed degree sequence and the target distribution for mixing between a pair of communities. To understand the ramifications of mixing between communities on effect size, we considered a range of mixing values. The networks are generated assuming random mixing by activity level, that is, the probability of forming a partnership does not depend on the number of partnerships an individual has. For each relationship, a date is drawn from a uniform distribution on the interval from start to end of study. The date is randomly chosen to represent the start or end of the relationship to avoid time trends in the number of relationships. The relationship duration is drawn from a duration distribution estimated from self-reported data in the Mochudi study. For relationships with durations longer than the study period, we calculate the likelihood of such relationships presenting during the entire study period based on the study duration and the relationship duration, and then adjusting either the start or end time to allow the possibility of such events accordingly. A histogram of the partnership durations is given in Figure 2.

[Figure 2 about here.]

Simulate the disease epidemic

In addition to data from the Mochudi study and the Botswana/Durban cohort, our model takes into account community characteristics including population size, varying coverage levels for different prevention modalities, as well as individual characteristics including transmission risk, disease progression, condom use, linkage to care, and circumcision status.

At time study time 0, we set the initial condition for each community. Each eligible individual is assigned an initial HIV infection status based on the current prevalence in Botswana, estimated to be 24.8%. Each infected individual is assigned to a viral load category as well as an initial CD4 count based on estimates of their distributions from the household survey in Mochudi. For CD4 counts below threshold for treatment, subjects are modeled as receiving ART according to estimates from Mochudi. The percentage of condom use is set as 40% and male circumcision rate at the start, at 12.7%, the estimated rate for Botswana(BAIS 2008). Background ART coverage for CD4<350 is set at 60.9% at the start based on a recent survey of the Mochudi district in Botswana in 2011. The probability of transmitting to a partner is based on the infected individual's viral load category, awareness of infection status, circumcision status, and treatment status, each of which is subject to change over time. For example, as disease progress, a subject's CD4 count may decrease to below threshold for treatment guidelines and making the subject eligible for treatment. The disease can only spread to a partner during the duration of the partnership. Disease progressions are assumed to follow estimates based on the Botswana/Durban cohort. Transmission risks to partners for different viral load categories are based on data reported in Quinn et al. (2000); additional sensitivity analyses are performed using rates reported in Attia et al. (2009) and Lingappa et al. (2010). Reductions in transmission risks associated with knowing infection status and condom use are set as 30% and 85%. Reduction in HIV acquisition risks for circumcision is set as 60%. To mimic the study design, we randomly pick 20% of population in each community to form the incidence cohort. Subjects in the incidence cohort are tested annually for HIV infection, and subjects outside of this cohort are tested with probabilities set to be the specified coverage levels for testing. The rates for circumcision, HIV testing and counseling and linkage to care (Table 1) are chosen to be the targeted levels for the intervention communities and the current levels for the standard of care communities. These coverage levels are allowed to vary yearly, therefore, the model allows assessment of the impact of a slower-than-expected intervention roll-out. In the standard of care commu-

nities, subjects become eligible for treatment based on national treatment guidelines; in the intervention communities subjects identified as high viral load carriers (>10,000 copies/ml) are also eligible for treatment.

[Table 1 about here.]

Effect of sexual mixing between communities

Sexual mixing between two communities arises when subjects in one community have partners from a different community. Mixing between intervention and standard- of-care communities would be expected to increase cumulative incidence in the former and decrease it in the latter, thereby attenuating the intervention effect. Simulating sexual networks with different levels of mixing permits assessment of the effect of mixing on treatment effect and subsequently on power and sample size estimation. Figure 3 illustrates the impact of increasing levels of mixing while holding other conditions fixed, the effect of which is to make the more similar the cumulative incidences in two sets of communities. When the mixing level reaches 50% – a level that implies that subjects are equally likely to have partners within and outside of their community – the expected cumulative incidence rates becomes very similar.

[Figure 3 about here.]

Self-reported data from the Mochudi study suggest that approximately 30% of partnerships were formed outside of that community. Mixing between communities randomized to the same intervention or between SOC communities and those not in the study does not attenuate intervention effects. Furthermore, many Mochudi residents work in the nearby capital city Gaborone, the residence of a considerable number of outside partners; by contrast most villages are relatively far from major urban centers. Therefore for our setting, we choose a lower level of mixing, 20%, with standard error 2.5%. These choices imply that about

95% of sampled values will be between 15% to 25%. Table 2 below presents the projected cumulative HIV incidences in SOC and intervention communities over 4 years of follow-up.

[Table 2 about here.]

Coefficient of variation

In addition to intervention effect, sample size calculation for a cluster randomized trial requires knowledge of coefficient of variation k, a parameter reflecting variations across communities induced by correlations among study subjects within the same community. From our simulations, we can obtain plausible values of k by dividing the sampling standard deviation of the cumulative incidences for the standard-of-care communities from many repeated experiments by their sample mean. In our model, all clusters are assumed to have the same population sizes, initial conditions, and rates of disease progression for infected subjects. These actually vary over communities, but because our study employs a matched-pair design, the appropriate k is the coefficient of variation in true proportion infected between matched communities in the absence of intervention. Matched pairs are intended to be quite similar in conditions, justifying the decision to keep them identical in simulations. Our simulation study yielded a k about 0.24. As some heterogeneity is likely to be present even across matched communities and has the effect of increasing k, we calculate power for values of k ranging from 0.2 to 0.35. Figure 4 displays the number of clusters and cluster sizes needed to achieve >90% power to detect the projected difference in 3-year cumulative incidences in SOC communities and intervention communities.

[Figure 4 about here.]

Fifteen clusters per arm and 500 incidence cohort members per community yields 96% power to detect the anticipated difference in cumulative HIV incidence between SOC and intervention communities (3.82% vs. 2.24%) by the end of the third study year, for $k = 0.24$ and

91% power for $k = 0.30$.

The sample size formula we used can be derived from random effects models in which cluster-level random effects are assumed to be independent across clusters, as are individual outcomes within clusters (e.g., an exchangeable correlation structure). We are concerned about deviations from the exchangeable-correlation assumption because it is likely to be violated for the outcome of HIV infection: for example, correlation between two individuals in a partnership would be expected to be higher than that between two individuals who are far apart in a sexual network but within a community. Fortunately, we found that deviations from an exchangeable correlation assumption do not affect the validity of the sample size formula. Under the exchangeable correlation structure, the intraclass correlation represents correlation between any 2 subjects in the same cluster. When the exchangeable correlation assumption is violated, the intraclass correlation ρ does not have this interpretation; instead, it reflects average correlation of observations from the same cluster. Even with arbitrary variance-covariance structure within cluster, however, the increase in variance resulting from cluster sampling, commonly measured by the design effect (Kish 1965), can be expressed by a function of the intraclass correlation ρ and the number of subjects within cluster. The between-cluster coefficient of variation k – which provides equivalent information regarding variance inflation as the intraclass correlation – captures the heterogeneity in outcomes across clusters resulting from the correlations among subjects from the same cluster.

To illustrate the assertions above, we consider the general setting where we have c clusters and sample m subjects within each cluster. Let Y_{ik} represent the outcome of the k^{th} ($k = 1, \ldots, m$) individual in i^{th} ($i = 1, \ldots, c$) cluster and we consider the following random effects model with correlated error terms:

$$Y_{ik} = \mu + \alpha_i + \epsilon_{ik},$$

where $\alpha_i \sim N(0, \sigma^2_{BC})$, $(\epsilon_{i1}, \ldots, \epsilon_{im})^T \sim V = (\sigma_{k,\ell})$, a variance-covariance matrix of dimension m. Let $\sigma^2_{WC} = \sigma_{kk}$ and $\sigma^2 = \sigma^2_{BC} + \sigma^2_{WC}$ denote the total variance. The correlation matrix for the m subjects within the same community is:

$$\left\{ \begin{array}{ccccccc} 1 & \rho_{1,2} + \rho_C & \cdots & \rho_{1,\ell} + \rho_C & \cdots & \cdots & \rho_{1,m} + \rho_C \\ \rho_{2,1} + \rho_C & 1 & \cdots & \rho_{2,\ell} + \rho_C & \cdots & \cdots & \rho_{2,m} + \rho_C \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{m-1,1} + \rho_C & \cdots & \cdots & \cdots & \cdots & 1 & \rho_{m-1,m} + \rho_C \\ \rho_{m,1} + \rho_C & \cdots & \cdots & \cdots & \cdots & \rho_{m,m-1} + \rho_C & 1 \end{array} \right\}$$

where $\rho_C = \frac{\sigma^2_{BC}}{\sigma^2}$, $\rho_{k,\ell} = \frac{\sigma_{k,\ell}}{\sigma^2}$.

The intraclass correlation ρ is defined as $\frac{E\{(Y_{jk}-\mu)(Y_{j\ell}-\mu)\}}{E(Y_{ik}-\mu)^2}$, where the expectation in the numerator is over all distinct pairs of individuals ($k \neq \ell$) taken from the same cluster and over all clusters and the expectation in the denominator is taken over all individuals and all clusters. In our setting, ρ can be expressed as

$$\frac{\sigma^2_{BC} + \{m(m-1)\}^{-1} \sum_{1 \leq k \neq \ell \leq m} \sigma_{k,\ell}}{\sigma^2}.$$

The corresponding between-cluster variance $\sigma^2_B = \sigma^2_{BC} + \{m(m-1)\}^{-1} \sum_{1 \leq k \neq \ell \leq m} \sigma_{k,\ell}$ and the coefficient of variation

$$k = \frac{\sigma^2_B}{\mu} = \frac{\sigma^2_{BC} + \{m(m-1)\}^{-1} \sum_{1 \leq k \neq \ell \leq m} \sigma_{k,\ell}}{\mu}.$$

Following the similar derivations for the design effect under an exchangeable correlation structure outlined in Hayes and Moulton (2009), it can be shown that the design effect ($DEFF$) in this case becomes:

$$DEFF = 1 + (m-1)\rho_C + \frac{\sum_{1 \leq k \neq \ell \leq m} \rho_{k,\ell}}{m} = 1 + (m-1)\rho.$$

To estimate between-cluster variance σ^2_B and coefficient of variance k, it is sufficient to use summary measures from each cluster. Let $\overline{Y}_{i\cdot}$ denote the individual cluster means, $\overline{Y}_{\cdot\cdot}$ denote

the overall mean, and s^2 denote the empirical variance of cluster means $\overline{Y}_{i.}$, we can estimate σ_B^2 and k as follows:

$$\widehat{\sigma}_B^2 = s^2 - \frac{\widehat{\sigma}_{WC}^2}{m}, \text{ and } \hat{k} = \frac{\widehat{\sigma}_B^2}{\overline{\overline{Y}}_{..}}.$$

Table 3 presents the actual and estimated k, ρ, and $DEFF$ under an arbitrary correlation structure and illustrates that the design effect estimated based on intraclass correlation ρ is unbiased. Here, we assume that there are 30 communities and each community has 20 individuals enrolled in the study. We let $\sigma_{WC} = 1$ and $\mu = 1$. As σ_{BC} increases from 0 to 1, the design effect increases from 1.5 to 10.8. The results are based on 1000 simulated studies. Although the true correlation matrix within a cluster is arbitrary, a single parameter ρ is sufficient to summarize the complex correlation structure and k captures the heterogeneity in outcomes across clusters induced by the complex correlation structure.

[Table 3 about here.]

We note that when departure from exchangeable correlation structure is expected, it is important that the studies used to estimate k employ the same sampling strategy as will the proposed study. Consider a special case where outcomes of individuals within the same households are more correlated than those of individuals from different households within the same community. Assume that the sampling strategy is such that within each of the c clusters, we randomly sample a households and b subjects within each household.

Using the similar notation as before, the true data generating process can be expressed as:

$$Y_{ijk} = \mu + \alpha_i + \gamma_j + \epsilon_{ijk},$$

where $i = 1, \ldots, c$ represents clusters, $j = 1, \ldots, a$ represents households, and $k = 1, \ldots, b$ represents subjects. We assume that $\alpha_i \sim N(0, \sigma_{BC}^2)$, $\gamma_j \sim N(0, \sigma_{BH}^2)$, $\epsilon_{ijk} \sim N(0, \sigma_{WH}^2)$.

Under this model,

$$var(Y_{ijk}) = \sigma^2_{BC} + \sigma^2_{BH} + \sigma^2_{WH} \stackrel{\Delta}{=} \sigma^2,\ cov(Y_{ijk}, Y_{ijk'}) = \sigma^2_{BC} + \sigma^2_{BH},\ cov(Y_{ijk}, Y_{ij'k'}) = \sigma^2_{BC}.$$

Let $\rho_H = \frac{\sigma^2_{BH}}{\sigma^2}$ and $\rho_C = \frac{\sigma^2_{BC}}{\sigma^2}$, the corresponding correlation matrix $(\rho_{k,\ell})$ within each cluster is

$$\left\{ \begin{array}{ccccccc} 1 & \rho_H + \rho_C & \cdots & \rho_H + \rho_C & \rho_C & \cdots & \rho_C \\ \rho_H + \rho_C & 1 & \cdots & \rho_H + \rho_C & \rho_C & \cdots & \rho_C \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_C & \cdots & \rho_C & \cdots & \cdots & 1 & \rho_H + \rho_C \\ \rho_C & \cdots & \rho_C & \cdots & \cdots & \rho_H + \rho_C & 1 \end{array} \right\},$$

that is, $\rho_{k,\ell} = \left\{ \begin{array}{ll} 1 & \text{if } k = \ell \\ \rho_C & \text{if } k \neq \ell, \text{and subjects } k \text{ and } \ell \text{ are in different household} \\ \rho_H + \rho_C & \text{if } k \neq \ell, \text{and subjects } k \text{ and } \ell \text{ are in the same household} \end{array} \right.$.

In the absence of household heterogeneity, $\sigma^2_{BH} = 0$ and $\rho_H = 0$, the correlation structure would be exchangeable. The current model allows that subjects in the same households have stronger correlation than subjects from different households in the same community.

The between-cluster variance $\sigma^2_B = \sigma^2_{BC} + \frac{b-1}{ab-1}\sigma^2_{BH}$. If we sample one person per household, $b = 1$, then the coefficient of variation $k_1 = \frac{\sigma^2_{BC}}{\mu}$; when we sample more than one people per household, $b > 1$, $k_2 = \frac{\sigma^2_{BC} + \frac{b-1}{ab-1}\sigma^2_{BH}}{\mu}$, and $k_1 < k_2$ for any $b > 1$. Therefore, if the previous study sampled one individual per household and the proposed study intended to sample entire households, the estimated k from the previous study would underestimate the true coefficient of variation for the proposed study and could result in an under-powered study.

Discussion

The properties of proposed study designed were investigated using simulation studies of epidemics propagated on collections of sexual networks. These studies allow us to take into consideration some important sexual network characteristics, such as mixing within and between communities, as well as coverage levels for different prevention modalities in the

combination prevention package under study.

Matching of villages is most effective when based on accurate predictions of incidence in the absence of intervention. Such predictions could be based on current estimates of incidence – for example obtained from using cross-sectional methods that rely on an incidence assay (e.g., the BED assay). In the proposed Botswana study, however, the anticipated number of eligible subjects in each community is too small (only about 670 on average) for this approach to be useful. Even based on optimistic assumptions (such as perfect assay specificity), only when the incidence rates for 2 communities differ greatly, say 1% versus 6-7% or higher, would we be able to distinguish the difference from random variation. Such high levels of incidence have not been observed in Botswana. Furthermore, use of assay-based incidence estimates for matching would require the study team to wait until all or most of the communities had completed the household surveys before any could be randomized and study intervention could be initiated. This could delay the start of the study by almost a year, it is preferable to randomize within pairs of communities and initiate the interventions as soon as the household surveys are completed in both members of each pair.

Mathematical modeling plays a critical role in planning and evaluating treatment for prevention (WHO report). As with any models, it is essential to understand the underlying assumptions and the impact of different choices of input parameters (Goyal et al., 2012). We construct our sexual network based on data from Likoma Island, and base disease progression for incident cases and prevalent cases on longitudinal estimates from the Botswana/Durban incidence cohort (n=77) and the Mochudi study. It would be desirable to have alternative network data to assess the robustness of our model results, as well as precise estimates based on larger sample sizes for disease progression. Nonetheless, our model estimates for the annual and cumulative incidence rates of the SOC communities are in line with the projected estimates from the UNAIDS Spectrum model (`http://www.unaids.org/en/dataanalysis/`

datatools/spectrumepp2013/). We also performed extensive sensitivity analyses for scenarios associated with lower-than-projected treatment effects (Figure 5). These analysis demonstrated that, for the planned sample size and k of 0.24, the study has >80% power to detect a reduction of 32% in the cumulative incidence in the intervention arm compared to the SOC arm (3.82%).

[Figure 5 about here.]

The proposed Botswana study is one of the two large HIV prevention trials commissioned by PEPFAR that are currently in development. The other is HPTN 071, the PopART study, which investigates a combination of interventions including universal testing, counseling and ART in Zambia and South Africa. The two trials differ in the components of combination prevention. A special feature of the Botswana study is its focus on identifying high viral carriers and treating them with ART. Both studies rely on extensive mathematical modeling to investigate the plausibility of different effect sizes of intervention. These models make use of information from a wide variety of sources regarding biology and behavior information that will be updated during the course of the studies and at their completion.

References

1. Attia S, Egger M, Muller M, Zwahlen M, and Low N. Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. *AIDS* 2009; 23:000-000.

2. BAIS. The Botswana HIV/AIDS impact survey III results, 2008. http://www.gov.bw/Global/NACA\%20Ministry/wana/BAIS\%20III_Stats\%20Press.pdf.

3. Blitzstein, J., and Diaconis, P. A Sequential Importance Sampling Algorithm for Generating Random Graphs with Prescribed Degrees, *Internet Mathematics* 2011; Vol. 6, Iss. 4, 2011

4. Boily M-C, Masse B, Alsallaq R, Padian NS, Eaton JW, et al. HIV treatment as prevention: considerations in the design, conduct, and analysis of cluster randomized controlled trials of combination HIV prevention. *PLoS Med* 2012; 9(7): e1001250. Doi: 10.1371/journal.pmed.1001250.

5. Botswana Central Statistics Office. Botswana population and housing census 2011. `http://www.cso.gov.bw/index.php?option=com_content1&id=2&site=census`.

6. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med* 2011; 365: 493-505.

7. Goyal R, Blitzstien J, and De Gruttola V. Estimation of degree mixing matrices with applications to network analysis and HIV prevention programs. *Harvard University Biostatistics Working Paper Series* 2011; `http://biostats.bepress.com/harvardbiostat/paper137`.

8. Goyal R, Wang R, De Gruttola V. Network epidemic models: assumptions and interpretations. *Clinical Infectious Diseases* 2012; 55(2):276-8.

9. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG. Universal voluntary testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *Lancet* 2009; 373: 48-57.

10. Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., and Morris, M. statnet: Software tools for the Statistical Modeling of Network Data., URL http://statnet.org, 2003.

11. Hayes RJ and Moulton LH. Cluster Randomized Trials. Chapman & Hall/CRC. Boca Raton, FL, USA, 2009.

12. Helleringer S and Kohler H. Sexual network structure and the spread of HIV in Africa: evidence from Likoma Island, Malawi. *AIDS* 2007; 21: 2323-2332.

13. Hughes J and Kulich M. Cluster randomized trials for HIV prevention. *Current Opinion in HIV & AIDS* 2006; 1(6): 471-475.

14. Jones JH and Handcock MS. An assessment of preferential attachment as a mechanism for human sexual network formation, *Proceedings of the Royal Society, B*, 2003; 270: 1123-1128.

15. Kish L. Survey Sampling, New York: John Wiley & Sons, Inc, 1965.

16. Lingappa JR, Hughes JP, Wang RS, Baeten JM, Celum C, Gray GE, et al. Estimating the impact of plasma HIV-1 RNA reductions on heterosexual HIV-1 transmission risk. *PLoS ONE* 2010; 5(9): e12598.

17. National AIDS Coordinating Agency (NACA). Analysis of HIV prevention response & modes of transmission. Gaborone: Government of Botswana, 2010.

18. Novitsky V, Wang R, Bussmann H, Lockman S, Baum M, et al. HIV-1 subtype C-infected individuals maintaining high viral load as potential targets for the test-and-treat approach to reduce HIV transmission. *PLoS ONE* 2010; 5(4): e10148. Doi:10.1371/journal.pone.0010148.

19. Novitsky V, Ndungu T, Wang R, Bussmann H, Chonco F, et al. Extended high viremics: a substantial fraction of individuals maintain high plasma viral RNA levels after acute HIV-1 subtype C infection. *AIDS* 2011; 25:1515-1522.

20. Novitsky V and Essex M. Using HIV viral load to guide treatment-for-prevention interventions. *Curr Opin HIV AIDS* 2012; 7:117-124.

21. Novitsky V, Woldegabriel E, Kebaabetswe L, Rossenkhan R, Mlotshwa B, Bonney C, et al. Viral load and CD4+ T cell dynamics in primary HIV-1 subtype C infection. *J Acquir Immune Defic Syndr* 2009; 50:65-76.

22. Novitsky V, Woldegabriel E, Wester C, McDonald E, Rossenkhan R, Ketunuti M, et al. Identification of primary HIV-1C infection in Botswana. *AIDS Care* 2008; 20:806-811.

23. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. *N Engl J Med* 2000; 342: 921-9.

24. World Health Organization, Meeting report on framework for metrics to support effective treatment as prevention, 2012.

25. Wright JK, Novitsky V, Brockman MA, Brumme ZL, Brumme CJ, Carlson JM, et al. Influence of Gag-protease-mediated replication capacity on disease progression in individuals recently infected with HIV-1 subtype C. *J Virol* 2011; 85:3996-4006.

Figure 1. An illustrative example of a static network of 2 communities. Open circles and solid circles represent individuals in different communities. Within each community, the location of circles does not represent their geographical locations.

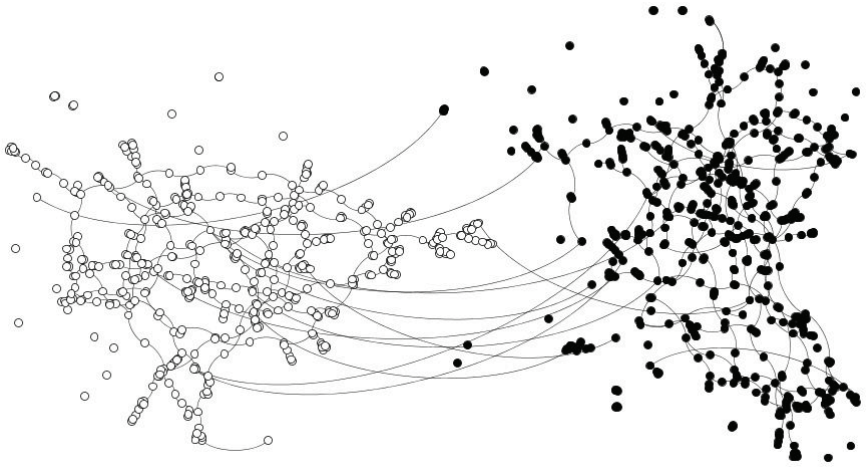

Figure 2. Relationship Durations in Mochudi

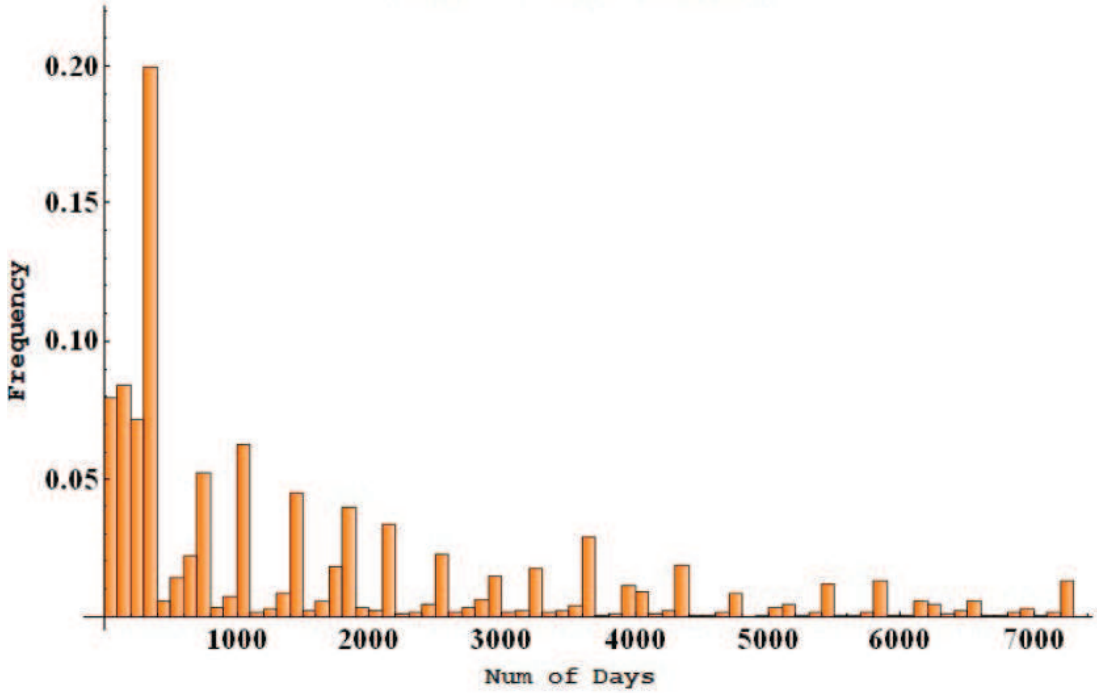

Figure 3. Cumulative incidence of intervention and standard-of-care communities over the 3-year period with varying levels of mixing

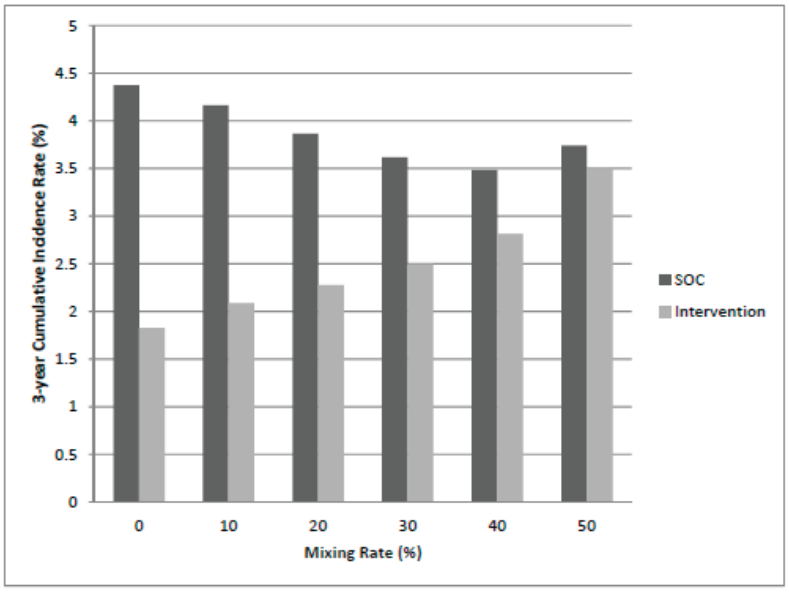

Figure 4. Number of Clusters versus Cluster Size Needed to Ensure >90% Power to Detect Anticipated Differences in Cumulative HIV Incidence Between SOC and Intervention Arms

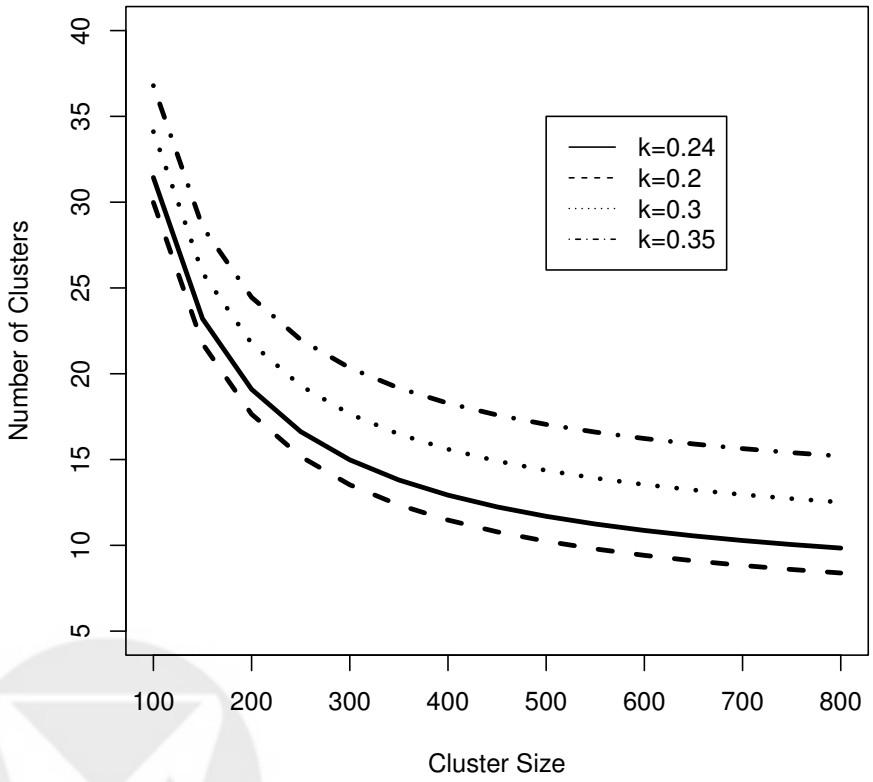

Figure 5. Power to detect various potential reductions in cumulative HIV incidence

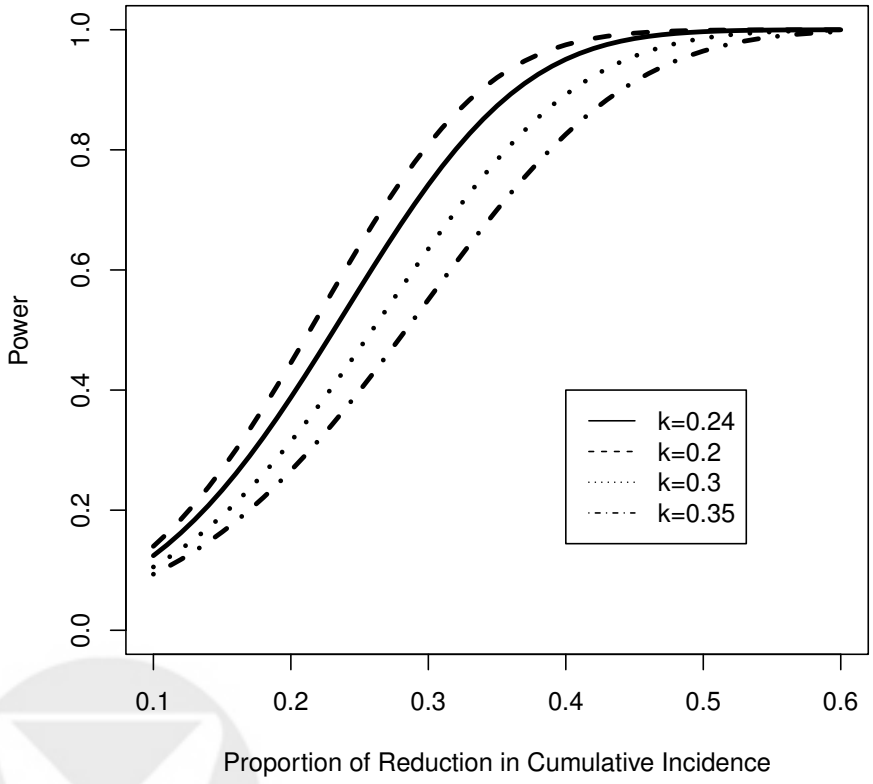

Table 1

Model input parameters to estimate impact of CP package scale-up in intervention communities versus SOC communities over 4 years

	SOC Arm			Intervention Arm		
	HTC	Male Circumcision	Linkage to Care	HTC	Male Circumcision	Linkage to Care
Baseline	37%[1]	12.7%[1]	80%	37%	12.7%[1]	90%
End Y1	37%	31.4%[2]	80%	81%	46.4%[4]	90%
End Y2	45%	50.0%[2]	80%	90%[3]	80%[4]	90%
End Y3	52%	60.0%[2]	80%	90%[3]	80%[4]	90%
End Y4	59%	70.0%[2]	80%	90%[3]	80%[4]	90%

1. BAIS 2008.
2. Male circumcision campaigns in SOC communities will be ongoing, and may reach 70% coverage by the end of Y4 post randomization, if Ministry of Health targets are met.
3. Assume that the project aims to increase HTC coverage to $\geq$90% in intervention communities by the end of the second study year and maintain this thereafter.
4. Assume that the project aims to reach 80% MC coverage in intervention communities by the end of the second study year and maintain this thereafter.

Table 2

Projected cumulative HIV incidence in SOC versus intervention communities over 4 years of study follow-up

	SOC	Intervention	
	Cumulative Incidence	Cumulative Incidence	% Reduction
End of Y1	1.71%	1.26%	26.3%
End of Y2	2.90%	1.84%	36.7%
End of Y3	3.82%	2.24%	41.4%
End of Y4	4.51%	2.58%	42.9%

Table 3

Actual and estimated coefficient of variation k, intraclass correlation ρ, and design effect. $\widehat{E}$ denotes average of estimates from 1000 simulated studies.

σ_{BC}	ρ	$\widehat{E}(\hat{\rho})$	k	$\widehat{E}(\hat{k})$	DEFF	$\widehat{E}(\widehat{DEFF})$
0.0	0.027	0.028	0.164	0.153	1.512	1.522
0.1	0.037	0.036	0.192	0.181	1.695	1.687
0.2	0.064	0.064	0.259	0.253	2.223	2.220
0.3	0.107	0.108	0.342	0.339	3.039	3.049
0.4	0.161	0.158	0.432	0.426	4.062	4.007
0.5	0.222	0.219	0.526	0.525	5.210	5.170
0.6	0.285	0.286	0.622	0.635	6.406	6.437
0.7	0.347	0.341	0.719	0.713	7.592	7.478
0.8	0.407	0.409	0.817	0.835	8.727	8.780
0.9	0.462	0.461	0.915	0.929	9.786	9.766
1.0	0.513	0.514	1.013	1.052	10.756	10.759