AIDS Behav (2013) 17:2893–2901
DOI 10.1007/s10461-013-0501-5

ORIGINAL PAPER

The Utility of a Composite Biological Endpoint in HIV/STI Prevention Trials

Tyler D. Hartwell · Willo Pequegnat · Janet L. Moore · Corette B. Parker · Lisa C. Strader · Annette M. Green · Thomas C. Quinn · Judith N. Wasserheit · Jeffrey D. Klausner · NIMH Collaborative HIV/STD Prevention Trial Group

Published online: 9 June 2013

Abstract A human immunodeficiency virus (HIV) as a biological endpoint in HIV prevention trials may not be feasible, so investigators have used surrogate biological outcomes. In a multisite trial, the epidemiology of STIs may be different across sites and preclude using one STI as the outcome. This study explored using a composite STI outcome to address that problem. The combined biological endpoint was the incidence of any of six new STIs (chlamydia, gonorrhea, trichomonas (women only), syphilis, herpes simplex virus type 2 infection and HIV) during a 24-month follow up period. We investigated how a composite STI outcome would perform compared to single and dual STI outcomes under various conditions. We simulated outcomes for four populations that represented a wide range of sex and age distributions, and STI prevalences. The simulations demonstrated that a combined biologic outcome was superior to single and dual STI outcomes in assessing intervention effects in 82 % of the cases. A composite biological outcome was effective in detecting intervention effects and might allow more investigations to incorporate multiple biological outcomes in the assessment of behavioral intervention trials for HIV prevention.

Keywords Sexual behavior · Sexually transmitted diseases · HIV/STI prevention · Simulation of biological outcomes

Introduction

Composite biological endpoints are increasingly being used in clinical trials to demonstrate the efficacy of an experimental intervention. However the way these composite endpoints are designed and interpreted has not been evaluated [1]. In human immunodeficiency virus (HIV) and sexually transmitted infection (STI) prevention trials, multiple biologic outcomes have been combined to increase the power of a study and increase the likelihood of detecting an intervention effect [2]. This strategy can maximize the possibility of obtaining an outcome for every participant, as well as a single primary outcome for the trial, even when there are missing data (i.e., one or more STI test results are missing for an individual but other STI test results are available) or individuals have entered the trial with incurable STIs (e.g., HIV and herpes simplex virus type 2 [HSV-2]). The most meaningful and appropriate measure of the efficacy of a behavioral HIV prevention intervention is a reduction in HIV incidence, but it is often not feasible to use HIV infection as a primary endpoint because of the low incidence of HIV seroconversion and the modest expected effect size (30–50 %) of

This study is conducted on behalf of the NIMH Collaborative HIV/STD Prevention Trial Group.

T. D. Hartwell · J. L. Moore · C. B. Parker · L. C. Strader · A. M. Green
Research Triangle Institute, Research Triangle Park, Durham, NC, USA

W. Pequegnat (✉)
National Institute of Mental Health, 6001 Executive Boulevard, Room 6111, Bethesda, MD 20892-9621, USA
e-mail: wpequegn@mail.nih.gov

T. C. Quinn
Johns Hopkins University, Baltimore, MD, USA

J. N. Wasserheit
University of Washington, Seattle, WA, USA

J. D. Klausner
University of California, Los Angeles, Los Angeles, CA, USA

behavioral interventions. Researchers have, therefore, used incident HIV infection and STIs to demonstrate efficacy of behavioral and biomedical interventions [3, 4]. Yet, those studies did not describe how the selection and use of multiple outcomes impacted the interpretation of the study findings.

Using a composite biological endpoint of STIs, including HIV infection, presents multiple challenges because of different biological and transmission characteristics such as infectivity, duration of infectiousness, symptomatology, ease of diagnosis and treatment, and risk of repeat infections [5]. Therefore, it is not clear that combining multiple biological outcomes with different epidemiologic characteristics will produce interpretable results; more may not be better.

The purpose of the study reported in this paper was to develop empirical evidence to evaluate whether a composite biological outcome performs better than a single or dual (HIV and HSV-2) biological outcome in evaluating HIV prevention interventions in multiple study sites where the study populations may have different STI prevalence, demographic, and behavioral characteristics. By simulating various combinations of STI outcomes and modifying both the strength of the intervention effect and the number of STIs in the outcome measured, we demonstrated how a composite test statistic might vary. This is the first empirical analysis of how the selection of STI outcome measures might impact study interpretation in various epidemiologic settings.

Background

The selection of which biological endpoint to use was an important question in the National Institute of Mental Health (NIMH) Collaborative HIV/sexually transmitted disease (STD) Prevention trial (hereafter called the C-POL trial) [6]. The data utilized for the hypothetical populations reported in this paper were based on the NIMH trial experience. This trial adapted a community-based intervention for international sites with a range of sexually active populations at risk for STIs, including HIV infection. The primary trial objective was to determine if community population members in the community popular opinion leader (C-POL) intervention venues showed greater reductions in sexual risk practices and lower HIV/STD incidence than those in the comparison venues.

The C-POL trial was a two arm, community-level, cluster randomized controlled trial that involved between 20 and 40 independent clusters or venues (discrete geographic community sites) in each country where population members at increased risk for HIV infection and STIs could be reached by community popular opinion leaders (C-POLs) who delivered HIV/STI risk reduction messages. The objective was to change social norms for safer sexual practices in the target populations. The design of the C-POL trial, using a composite biological endpoint, and the results have been reported elsewhere [6, 7]. Using simulation techniques informed by the experiences from the C-POL trial, this study examines the performance of a composite biological endpoint to detect the effect of an intervention under various conditions. Based upon the observational data from the C-POL trial, we assumed the prevalence of the STIs varied widely at baseline across the populations.

Methods

Setting

The setting for this simulated analysis was a multi-population, two arm, cluster randomized trial of an intervention to reduce sexual risk behavior. The study design included intervention and comparison venues in each of several populations to demonstrate an intervention effect (i.e., a reduction in incidence) as determined by a composite biological endpoint. Six STIs were identified that could be analyzed using blood, urine, or vaginal samples: chlamydia (Ct), gonorrhea (Ng), trichomoniasis (Tv) in women only, syphilis (SYP), genital herpes (HSV-2), and HIV. The primary simulation question was to ascertain a reasonable biological endpoint with sufficient power to demonstrate an intervention effect in each population under various scenarios: (1) the incidence of all STIs are reduced in the intervention venues, (2) the incidence of one STI is reduced in the intervention venues, and (3) the incidence of STIs are reduced in both intervention and comparison venues, etc., when the prevalence of those STIs varied greatly by population. In particular, given the above context, we wanted to examine how a composite biological endpoint would perform in detecting an intervention effect compared to using individual STIs.

Composite Biological Endpoint

The composite biological endpoint was defined as the incidence of any new STI: Ct, Ng, Tv (women), HSV-2, SYP, or HIV observed during a 24-month follow-up period. Thus, for each participant, a composite binary variable was used to indicate whether or not a new case of at least one of the six STIs was detected during follow-up. Individuals with Ct, Ng, Tv, and SYP were assumed to be treated and cured following a positive test result at baseline and thus eligible for incident infection at follow-up. Since HIV and HSV-2 infections cannot be cured, an individual was classified as positive for a new case of any of the six infections during follow-up if there was any new positive test for Ct, Ng, Tv, SYP, HSV-2 (if negative at baseline), or HIV (if negative at

baseline). Otherwise, an individual was classified as negative for the composite outcome denoted as "any STI."

Simulation

Simulation is a means of studying outcomes that replicate real world experience. By using well-specified steps, researchers develop evidence that elucidates different simulated outcomes [8–11]. In this study, simulation was used to demonstrate how a composite biological outcome performed compared to a single or dual (any viral STI) biological outcome in detecting the effect of a behavioral intervention.

In conducting the simulation study for this paper, we performed the following steps:

Step one. Established four hypothetical populations with 1,000 eligible participants each in 20–40 venues (a discrete geographical area where participants assemble (e.g., bar, office, school)) for the four populations (see sample size calculations for number of venues below). The four hypothetical populations were assumed to be the entire eligible populations congregating in the selected venues (see Table 1 for the sample sizes in the hypothetical populations).

Step two. Based on the data from the C-POL trial, the eligible participants from the hypothetical population were randomly assigned specific characteristics: (1) age and gender, (2) then, based on their age and gender, randomly assigned as an individual having one or more STIs or not having one or more STIs at baseline and during the two-year follow-up period, and (3) if, over time, they were lost to follow-up. As in the C-POL trial, these participant characteristics varied from population to population with a wide range of values so that the results of the simulation would be generalizable to other populations of similar age and gender distributions.

Further, we assumed intraclass correlations (ICCs) among the venues (i.e., among venue variances) in each of the four populations for the prevalence of each of the six STIs between 0.005 and 0.02 as well as correlations among the prevalence of the six STIs [6]. The incidences took into account whether a participant had an STI at baseline. We assumed for Ct, Ng, Tv, and SYP that the STI was treated and cured at baseline, while for those infected with HIV and HSV-2 at baseline, infection was lifelong. Intraclass correlations between venues were also assumed for the incidence data, as well as correlations among the incidences of the six STIs. (Note that the ICCs and correlations were taken into account when determining if a participant had an STI or not.) Table 1 presents the samples sizes chosen as well as the age and gender distributions for each of the hypothetical populations. Table 2 presents overall STI prevalence at baseline and STI incidences over a two year period in the hypothetical population comparison venues that were used to assign whether an individual had any of the STIs or not at baseline or during the two-year follow-up period.

We used lost to follow-up rates in each venue of the hypothetical populations based on the age-gender distribution in each country (Males < 25 years of age lost to follow-up rate = 20 %; Females < 25 years, rate $= 15$ %; Males > 25 years, rate $= 15$ %; Females > 25 years, rate $= 10$ %). Then, we randomly assigned individuals as lost to follow-up in each venue from the hypothetical population based on their age and gender. As described above, the assigned ICCs, correlations, prevalence, incidences, and follow-up rates were all based on the observational data in the C-POL trial.

Step three. Blocked venues within the hypothetical populations into pairs based on "any STI" rates at baseline (the two venues with the highest baseline "any STI" rates in the first pair, etc.). Blocks were used to reduce variability and to maximize statistical power. For each matched pair of venues, we randomly assigned one to intervention (I) and one to comparison (C).

Step four. A random sample of participants was selected for each venue within each of the four hypothetical populations. To determine how large a sample to draw from

Table 1 Sample sizes within gender and age (in years) distributions for the four hypothetical populations

Population	Males <25 years *N* (%)	Males >25 years *N* (%)	Females <25 years *N* (%)	Females >25 years *N* (%)	Total[a] *N* (%)
A	12,390 (41.3)[b]	3,510 (11.7)	10,620 (35.4)	3,480 (11.6)	30,000 (100)
B	11,280 (47.0)	5,520 (23.0)	3,600 (15.0)	3,600 (15.0)	24,000 (100)
C	5,400 (18.0)	15,600 (52.0)	3,300 (11.0)	5,700 (19.0)	30,000 (100)
D	2,960 (7.4)	15,760 (39.4)	1,840 (4.6)	19,440 (48.6)	40,000 (100)

[a] Total hypothetical Population size over venues (venues in Population A = 30, in Population B = 24, in Population C = 30 and in Population D = 40)

[b] The size of hypothetical Population A in the category males <25 years is 12,390. The percentage of males <25 years is 41.3 % in hypothetical Population A (total Population A size is 30,000)

Table 2 Overall population STI prevalences at baseline and incidences over two years by STI for the four hypothetical populations

STI	Population A Prevalence (Incidence) (%)	Population B Prevalence (Incidence) (%)	Population C Prevalence (Incidence) (%)	Population D Prevalence (Incidence) (%)
Ct	1 (1)	5 (6)	6 (5)	10 (6)
Ng	1 (1)	2 (2)	3 (3)	4 (2)
HSV-2	30 (15)	20 (5)	15 (5)	10 (4)
Tv	20 (5)	15 (18)	9 (9)	4 (5)
SYP	1 (1)	7 (2)	5 (2)	2 (1)
HIV	15 (9)	8 (3)	5 (2)	1 (1)
"Any STI"	37 (20)[a]	34 (19)	26 (13)	25 (12)

Note Incidence in table are for comparison venues (Incidence for intervention venues are changed (i.e. lowered) in simulations to indicate intervention effects over time)

[a] The "any STI" statistic shows that in hypothetical Population A the prevalence of this statistic is 37 % and it's incidence is 20 %. Thus, Populations A and B have relatively high prevalences and incidences while Populations C and D have relatively low prevalences and incidences

each venue, we performed sample size calculations for each population (i.e., number of venues, number of participants per venue) based on: (1) an assumed 30 % effect size in the intervention venues; (2) ICCs observed in the C-POL trial; and (3) the incidences assumed in the comparison venues for the endpoint "any STI" [7, 12]. The sample sizes assumed 80 % power and a two-sided significance level of 0.05. (Note: in Population C, the number of venues selected in the simulation was fewer than needed to obtain 80 % power so that we might observe the results when a sampled population was underpowered.) Based on those assumptions, the following sample sizes were calculated (number of venues, number of participants sampled in each venue, total number of sampled participants for each random sample): (1) Population A—number of venues: 30, participants sampled in each venue: 125, total number of sampled participants: 3,750, (2) Population B—number of venues: 24, participants sampled in each venue : 150, total number of sampled participants: 3,600, (3) Population C—number of venues: 30, participants sampled in each venue: 150, total number of sampled participants: 4,500, (4) Population D—number of venues: 40, participants sampled in each venue: 100, total number of sampled participants: 4,000. Thus, for example in Population A, a random sample of 125 participants was selected from the 1,000 population members in the hypothetical population in each of the 30 venues—3,750 participants (Fig. 1).

Step five. The statistic "any STI" as well as other statistics (HIV, HSV-2) were computed for each participant in the selected samples from Step four. We also computed the statistic "any viral STI" which was created to indicate participants who tested newly positive for HSV-2 or HIV over the 2-year follow-up period. If a participant was lost to follow-up, the statistics were missing for that participant (i.e., none of this participant's data was used to compute the statistics "any STI," etc.,). No imputation was performed for those lost to follow-up participants.

Step six. The incidence rates for all sampled participants in each venue in each population who had "any STI," HIV, HSV-2 or "any viral STI" over the 2 years was then computed. The overall test statistic for the sampled participants within a population was taken as the average of the differences of the intervention minus the comparison incidences, across venue pairs with equal weight to each pair (e.g., in Population A the 15 incidence differences (where 15 was the number of intervention and comparison venue pairs) were computed and then averaged to obtain the overall test statistic) [6] (Fig. 1).

Step seven. A permutation test was used to determine the significance among test statistics for each population. This yielded a p value for each test statistic in each population. The permutation test is based on the randomization of venues within venue pairs to the intervention or comparison condition. Statistical significance was computed by considering all possible values each test statistic could have taken by permuting the random assignment of venues within venue pairs. Under the null hypothesis of no difference between intervention and comparison conditions, the statistical significance of the observed results was taken as the rank of the observed statistic among the possible permutations; p values were two-sided [13].

Step eight. Steps four to seven were then repeated 1,000 times (e.g., 1,000 different samples were drawn randomly from the venues in the four hypothetical populations). This yielded 1,000 p values in each hypothetical population for each test statistic so that the statistics could be compared (Fig. 1 describes the simulation process for Population A (i.e., the population could be different geographically or politically defined areas, such as country, city, school, etc.)).

Step nine. Steps four to eight were then repeated to observe how the test statistic performed (e.g., the number of times the test statistics detected a difference between intervention and comparison venues) under various

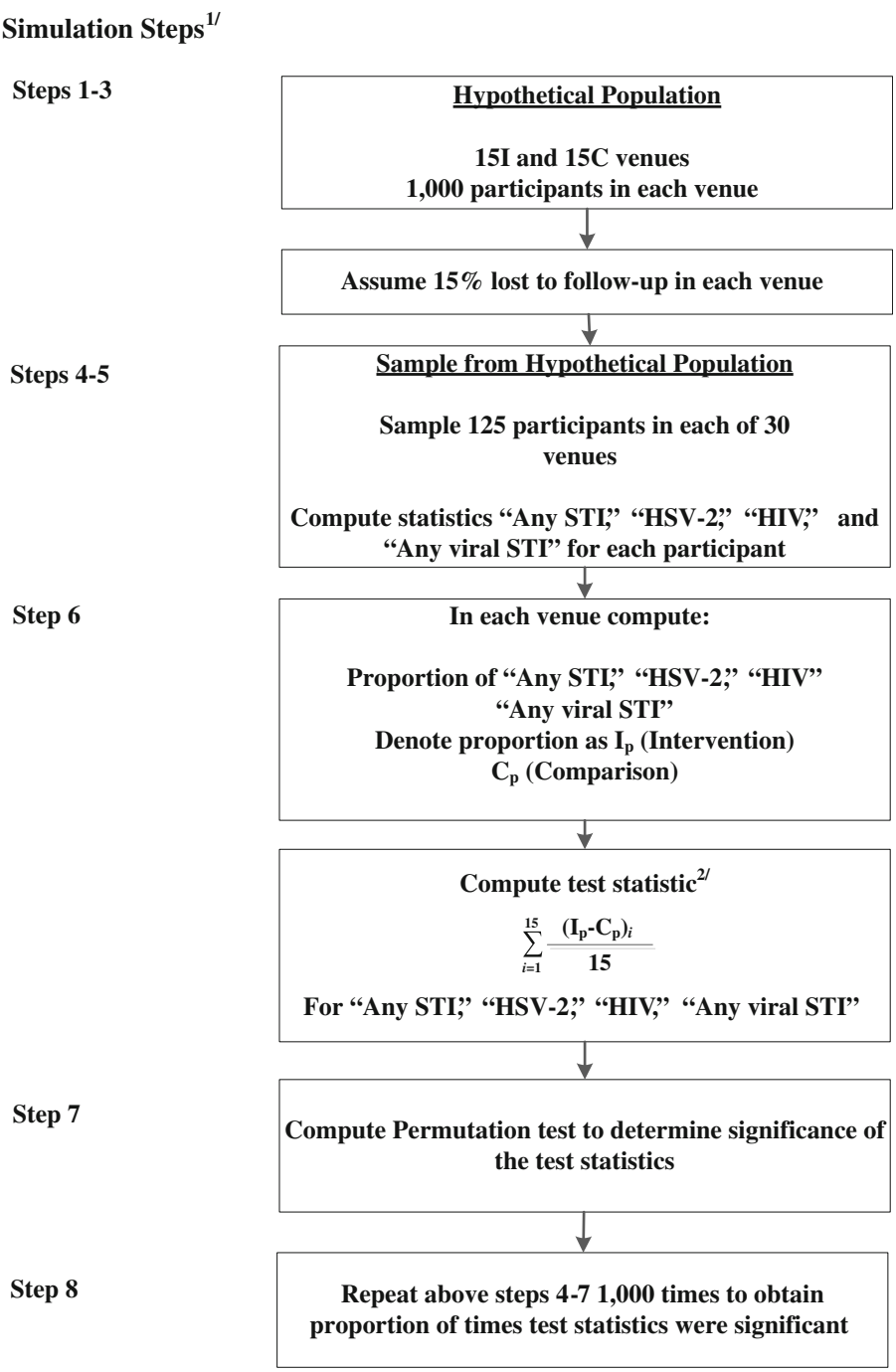

Fig. 1 Schematic of simulation process in country A

[1/] Described in Simulation Section of paper.

[2/] (I_p-C_p) difference in proportions in a matched pair of venues.

scenarios (e.g., a decrease of 30 % in incidence of all six STI's in the intervention venues compared to no change in incidence in the Comparison venues). The scenarios selected represented a wide range of possible intervention effects (e.g., STI incidence outcomes).

These steps allowed us to compare how each test statistic performed in detecting the difference between the intervention and comparison venues in each population assuming different effect sizes for the intervention. For example, for a particular intervention effect if the "any STI" p value was <0.05 (i.e., significant) in 900 cases out of 1,000 and an individual STI p value was <0.05 in 700 cases out of 1,000, then the "any STI" statistic was significant 90 % of the time compared to 70 % of the time for the individual STI.

The scenarios selected to compare the test statistics were:

1 A decrease of 30 % in the incidence of all six STIs in the intervention venues compared to the incidences in the comparison venues given in Table 2.
2 A decrease of 20 % in the incidence of all six STIs in the intervention venues compared to the incidences in the comparison venues given in Table 2.
3 A decrease of 30 % in the incidence of all six STIs in the intervention venues compared to a 10 % decrease in incidence of all six STIs in the comparison venues.
4 A decrease of 30 % in the incidence of HSV-2 only in the intervention venues compared to the incidences in the comparison venues given in Table 2.
5 A decrease of 30 % in the incidence of "any viral STI" only in the intervention venues compared to the incidences in the comparison venues given in Table 2.

These five scenarios were selected to compare the different test statistics under different conditions that represent a wide range of possible STI incident outcomes. Scenario 3 was designed to examine the case where there were decreases in incidence in both the intervention and comparison venues over the two-year intervention period, which is often the case in behavioral intervention trials.

Results

Table 3 summarizes the results of the 1,000 samples drawn from each of the four hypothetical populations for various intervention effect scenarios. In Table 3 the test statistic "any STI" is compared with the test statistics HSV-2, HIV and "any viral STI." HSV-2 was selected as one of the six STIs to study since it had the highest prevalence of any of the six STIs at baseline in the hypothetical populations (i.e., HSV-2 prevalence varied from 30 % in Population A to 10 % in Population D, Table 2).

The percentage of times that a difference between intervention and comparison venues, given various intervention scenarios, was almost always higher for the "any STI" statistic. For example, in Population A if the incidence of all six STIs was reduced by 30 % over the 2-year follow-up period then the "any STI" statistic detected this difference (i.e., p value <0.05) 99.6 % of the time while the statistics HSV-2, HIV and "any viral STI" detected it only 71.9, 64.8 and 93.4 % of the time, respectively. Similarly, in Population B if the incidence of all STIs were reduced by 20 % over the 2-year follow-up period then the "any STI" statistic detected the difference 97.9 % of the time, while the statistics HSV-2, HIV and "any viral STI" detected it only 7.8, 15.5 and 15.2 % of the time, respectively. Even in Population C where the number of venues (i.e., 30) was not adequate to pick-up reductions in the intervention venues 80 % of the time (recall that we purposely underpowered Population C), the "any STI" statistic detected the effect a higher percentage of time than the other statistics for scenarios 1, 2, and 3. In general, in populations where the STI prevalence and incidence rates were relatively high (i.e., Populations A and B), the "any STI" statistic was superior for all scenarios in Table 3. In countries where the STI rates were relatively low (i.e., Populations C and D), the "any STI" statistic was superior for scenarios 1, 2, and 3, but not for scenarios 4 and 5 in which detection differences were <50 % for all statistics.

Discussion

Based on observational data from a large multi-country behavioral intervention trial, we simulated the performance of a composite biological outcome versus single and dual biological outcomes. In most instances the composite biological outcome detected an intervention effect more often when compared with a single or dual outcome. We have shown how a composite STI outcome can contribute evidence towards determining the efficacy of an HIV/STI prevention intervention across multiple sites with different characteristics (e.g., STI prevalences and incidences).

In many settings, there are advantages of using a composite STI outcome:

(1) The composite STI endpoint has more power than individual STIs to detect intervention effects. This permits a reasonable sample size both within and across populations and a shorter follow-up period, which contributes to study feasibility. Intervention effects for individual STIs can also be examined as secondary endpoints but power may be limited in these cases.
(2) The composite STI endpoint allows the use of more STI data and participants within a study population; that is, if there is missing data on only one or two STIs for a participant, that participant can still be included in the outcome analysis. This is important because there may be varying degrees of missing data across STIs and the composite statistic permits researchers to measure an impact on the whole community (not only a subset as does a single STI). Also, the primary biological endpoint can be used across populations even though the populations might have different patterns of prevalence for the various STIs. Thus, there is one primary outcome rather than several outcomes as with individual STIs that can contribute to Type I errors.

Table 3 Percentage of time a test statistic detected a significant difference between intervention (I) and comparison (C) venues for different intervention effect scenarios and test statistics (any STI versus any viral STI, HSV-2, and HIV) by population

Intervention incidence effect scenarios over 2 years	Population															
	A				B				C				D			
	Any STI (%)	Any Viral STI (%)	HSV-2 (%)	HIV (%)	Any STI (%)	Any Viral STI (%)	HSV-2 (%)	HIV (%)	Any STI (%)	Any Viral STI (%)	HSV-2 (%)	HIV (%)	Any STI (%)	Any Viral STI (%)	HSV-2 (%)	HIV (%)
1. Incidence of all 6 STIs decrease by 30 % in I venues[a, b]	99.6[c]	93.4	71.9	64.8[c]	99.9	63.3	47.8	29.8	69.0	38.9	32.3	12.7	87.5	39.2	31.7	9.8
2. Incidence of all 6 STIs decrease by 20 % in I venues[a]	90.7	64.0	40.2	8.1	97.9	15.2	7.8	15.5	14.4	14.6	6.9	4.8	66.3	18.5	14.1	5.3
3. Incidence of all 6 STIs decrease by 30 % in I venues and 10 % in C venues	84.1	56.6	17.1	31.0	48.0	17.6	8.4	11.7	32.8	31.0	26.6	9.1	71.3	15.4	10.2	5.6
4. Only incidence of HSV-2 decreases by 30 % in I venues[a]	61.4	36.3	70.3	0.1	65.5	23.8	42.8	3.6	5.4	25.2	33.4	2.9	8.7	30.1	34.1	3.0
5. Only incidence of: "Any viral STI" decreases by 30 % in I venues[a]	91.5	84.3	71.6	69.7	82.0	36.9	24.1	17.9	9.8	38.6	37.2	12.3	14.4	38.7	29.6	9.6

[a] Assumed the incidence in comparison venues did not decrease from the values given in Table 2

[b] In this scenario the incidences of the STIs in the intervention (I) venues all decreased by 30 % from the incidences shown in Table 2 for the comparison (C) venues

[c] The "any STI" statistic in scenario 1 detected a significant difference between the I and C venues in Population A 99.6 % of the time while the HIV statistic only detected this difference 64.8 % of the time

(3) The composite STI endpoint also permits the same biological endpoint (a single composite measure) to be used for both men and women even though they may have different STIs and different collection methods.
(4) Results are more generalizable and may be more robust if several STIs are included in a biological endpoint rather than a single STI.
(5) Because the specific impact of a behavioral risk reduction intervention might not be known and the intervention may affect a variety of different sexual risk behaviors—type of sex, partner frequency, concurrent partnerships, condom use, etc.,—a composite STI endpoint provides for a more inclusive measure to determine possible impact.
(6) The cost of a study and its outcome measure(s) are always critical factors in study design. Because specimen collection for most STIs is similar (blood or genital secretions) and the cost of specimen testing is incremental (multiplex testing may be used to conduct multiple tests on the same clinical sample) several STIs may be used for a study without substantially increasing study cost. The treatment costs for most STIs like gonorrhea, chlamydia and trichomoniasis are single dose generic antibiotics in the range of $1–$5 per treatment. The treatment of STIs is a public health priority in many under resourced countries because they recognize their role in HIV transmission and complications in childbirth.

The strengths of a composite biological endpoint may outweigh several possible limitations including:

(1) Large-scale intervention trials focused on reducing HIV infection by treating STIs have delivered mixed results, reinforcing the supposition that biologic and epidemiologic relationships between STIs and HIV infection are complex [4]. The evidence provided by a composite STI endpoint depends on the pathogenesis, treatment, and transmission dynamics of specific STIs or HIV infection, the behavioral characteristics of the population, and the stage of the HIV epidemic. Therefore, there may be concerns regarding external validity when using a composite biological endpoint that incorporates different types of STIs (e.g., bacterial, protozoan, and viral).
(2) A composite STI outcome may not reflect the role of condom use in preventing the transmission of some STIs and not others, as well as the sexual behavior of the individuals. When condoms are used correctly and consistently, they are effective in reducing the incidence of HIV and some STIs (e.g., gonorrhea and chlamydia), but they may not be as effective in decreasing HSV-2 transmission, if viral shedding occurs in areas not covered by the condom [14].
(3) The simulations in this study were based on data from a sample of men and women recruited for a study in four populations using selection criteria based on high prevalence of HIV-related risk behaviors and STIs in the population. While those data may not be as representative as a country-wide random sample, they are representative of the high-risk populations outside the United States that would be recruited for HIV/STI clinical trials.

We believe that the results of this simulation study demonstrate that it is valid to use a composite biological outcome in behavioral intervention trials designed to reduce HIV infection, which strengthens the quality and interpretability of HIV/STI prevention research because the outcomes are based on more participant data (i.e., participants are not lost if they have one or two missing STI outcomes) in appropriately designed trials. Even sites that have different patterns of STIs can be included in a multisite study. Furthermore, the biological outcome data from men and women can be used together. Having both behavioral and biological outcomes provides complementary information on the efficacy of a program. Our simulation study demonstrated how a composite biological outcome might vary in different epidemiologic settings and was better than a single or dual biological outcome. These findings should reassure investigators that are already incorporating composite biological outcomes in behavioral intervention trials for HIV/STI prevention and encourage other investigators to adopt these methods.

Acknowledgment This study was funded by the National Institute of Mental Health.

References

1. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. BMJ. 2010;341:c3920.
2. Kamb ML, Fishbein M, Douglas JM Jr, et al. Efficacy of risk-reduction counseling to prevent human immunodeficiency virus and sexually transmitted diseases: a randomized controlled trial. Project RESPECT Study Group. JAMA. 1998;280(13):1161–7167.
3. Fishbein M, Pequegnat W. Evaluating AIDS prevention interventions using behavioral and biological outcome measures. Sex Trans Dis. 2000;27(2):101–10.
4. Pequegnat W, Fishbein M, Celentano D, et al. NIMH/APPC workgroup on behavioral and biological outcomes in HIV/STD prevention studies: a position statement. Sex Trans Dis. 2000;27(3):127–33.
5. NIMH Collaborative HIV/STD Prevention Trial. Results of the NIMH collaborative HIV/sexually transmitted disease Prevention

Trial of a community popular opinion leader intervention. J Acquir Immune Defic Syndr. 2010;24(2):204–14.
6. Garnett G. The transmission dynamics of sexually transmitted infections. In: Holmes KK, Sparling PF, Stamm WE, et al., editors. Sexually transmitted diseases. 4th ed. New York: McGraw-Hill; 2007. p. 27–40.
7. NIMH Collaborative HIV/STD Prevention Trial group. Methodological overview of a five-country community-level HIV/sexually transmitted disease Prevention Trial. AIDS. 2007;21(suppl 2): S3–18.
8. Alfons A, Kraft S, Templ M, Filzmoser P. Simulation of close-to-reality population data for household surveys with application to EU-SILC. Stat Methods Appt. 2011;20(3):383–407.
9. Goldman RN, McKenzie JD Jr. Creating realistic data sets with specified properties via simulation. Teach Stat. 2009;31(1):7–11.
10. Hagmark PE. On construction and simulation of count data models. Math Comput Simul. 2008;77(1):72–80.
11. Rubinstein RY, Kroese DP, Botev ZI, Taimre T. Simulation and the Monte Carlo method (Wiley series in probability and statistics, #707). 2nd ed. Hoboken: Wiley; 2007.
12. Murray DM, Hannan PJ. Planning for the appropriate analysis in school-based drug-use prevention studies. J Consult Clin Psychol. 1990;58(4):458–68.
13. Gail MH, Mark SD, Carroll RJ, Green SB, Pee D. On design considerations and randomization-based inference for community interventions trials. Stat Med. 1996;15(11):1069–92.
14. Holmes KK, Levine R, Weaver M. Effectiveness of condoms in preventing sexually transmitted infections. Bull World Health Organ. 2004;82(6):454–61.