Challenges and processes of selecting outcome measures for the NIMH Collaborative HIV/STD Prevention Trial

NIMH Collaborative HIV/STD Prevention Trial Group*

Objective: To review the challenges of designing behavioral and biological outcome measures for the multinational NIMH Collaborative HIV/STD Prevention Trial and provide the rationale for selecting these measures.

Design: Although many different evidence-based prevention programmes have been developed, few have been evaluated in different countries, cultures, and populations. One issue in evaluating the generalized efficacy of any prevention approach is to identify a set of common outcome measures useful across diverse settings and peoples. The Trial is designed to evaluate whether the community popular opinion leader intervention can be adapted cross-nationally and cross-culturally for different populations and still retain its efficacy.

Methods: Literature reviews, investigator experience, ethnographic study, pilot studies, and epidemiological studies were used to select the endpoints for the Trial.

Results and conclusion: Both biological and behavioral data will be obtained at baseline and 12 and 24 months post-baseline. Communities that receive the intervention will be compared with matched control communities on two primary outcomes: (i) a change in self-reported unprotected sexual acts with non-spousal, non-live-in partners; and (ii) the incidence of sexually transmitted disease (STD), defined as a composite index of viral and bacterial STD.

AIDS 2007, **21** (suppl 2):S29–S36

Keywords: AIDS, behavioral endpoints, behavioral intervention, biological endpoints, community popular opinion leader, developing nations, HIV, outcome measures, risk behavior

Introduction

Identifying effective behavioral and social science interventions to reduce the incidence of sexually transmitted disease (STD) and HIV has taken on a new urgency in response to the global HIV pandemic. Although a number of HIV prevention programmes have been shown to be efficacious in well-defined risk populations in the United States, these programmes must be adapted to new cultures and tested before they can be implemented where they are desperately needed [1]. A significant barrier to effectively demonstrating intervention efficacy is the lack of evidence-based measures of reductions in sexual risk behaviors. In designing interventions that can be used across diverse settings, prevention researchers need to develop relevant, measurable, and common primary outcomes applicable across countries and risk populations to demonstrate to public health leaders and policymakers the success or failure of the behavioral intervention to promote behavior change and reduce the incidence of disease.

Evaluating the impact of a behavioral intervention to prevent the acquisition of STD, including HIV, requires measuring sexual risk behaviors targeted by the intervention. Because sexual risk behaviors are part of the STD acquisition causal pathway, a change in those behaviors needs to occur to decrease the subsequent acquisition of infection [2]. Collecting reliable information about sexual risk behaviors may, however, be difficult because this is a private, taboo subject. Whereas some researchers argue that self-reported behavior can provide valid and reliable outcome measures, others worry that participant self-reports do not always accurately reveal sensitive behaviors [3,4]. The measurement of sexual risk behaviors alone is thus often considered insufficient in the evaluation of interventions to prevent STD.

To impact an epidemic, an intervention needs to decrease the potential for acquisition of a disease. Whereas risky HIV-related behavior is linked to acquiring STD, the correspondence is not one-to-one [2,3]. Transmission models conceptualize the reproductive rate of an

* See Appendix B for members of the NIMH Collaborative HIV/STD Prevention Trial Group.

infectious disease as the product of three components: (i) the rate of contact between infected and susceptible individuals; (ii) the transmissibility of infection between partners; and (iii) the duration of infectiousness [5–7]. Each of these components can be influenced by behavioral factors, such as the number of sex partners in a given period, the correct use of condoms for every sex act, and the type and duration of every sex act, as well as by environmental factors, such as the extent and accessibility of population-level screenings and effective treatments.

Whether a reduction in risk behaviors results in a reduction in the incidence of infectious diseases depends on the epidemiological context in which the intervention is implemented [8]. Grassly *et al*. [8] proposed four indicators describing the epidemiological context that should be considered when evaluating the appropriateness of an intervention to prevent HIV infection: (i) the phase of the HIV epidemic; (ii) the co-occurrence of other STD; (iii) the mixing of the target population with other at-risk populations; and (iv) the sexual behavior of populations not targeted by the intervention. The relationship between risk behaviors and biological outcomes is thus complex, non-linear, and impacted by multiple social, cultural, and environmental factors [3,4,9,10].

The US National Institute of Mental Health (NIMH) Collaborative HIV/STD Prevention Trial (hereafter 'the Trial') was initiated to evaluate whether the popular opinion leader intervention [11] can be adapted in multiple countries with various vulnerable, at-risk populations and retain its efficacy. The community popular opinion leader (C-POL) intervention being evaluated in the Trial is the international adaptation of the popular opinion leader intervention, which was found to be effective in reducing behavioral risk among homosexual men, adolescents, and heterosexual women in housing developments in the United States [12–16].

The Trial is being conducted in five countries: China, India, Peru, Russia, and Zimbabwe. Using data from preliminary studies, vulnerable subpopulations were identified and targeted for the intervention in each country: food market stall owners and workers in China, male patrons of wine shops and at-risk women congregating near the shops in India, young men and women in social gathering points in neighborhoods in Peru, trade and vocational school dormitory residents in Russia, and individuals congregating in growth points in Zimbabwe. Whereas the age range of eligible participants in each country was based on the HIV/STD epidemiology in the country, all countries included a core group of young adults aged 18–30 years who are likely to be among the most sexually active and vulnerable in the population.

Given the complex dynamics between behavior and changes in HIV/STD incidence, the Trial researchers identified the need for both behavioral and biological indices to evaluate the efficacy of this intervention. Experts in behavioral interventions, STD transmission, behavioral assessment, and rapid ethnography formed workgroups that participated in designing different aspects of the Trial (see 'Methodological overview of a five-country community-level HIV/sexually transmitted disease prevention trial' in this issue [17]). Intense review and debate occurred based on the results of both rapid ethnography and preliminary epidemiological studies of risky populations in each country that provided both behavioral and biological data. From this process, the two primary outcome measures for this collaborative community-randomized Trial were defined.

This paper presents the challenges and process of selecting an appropriate primary behavioral outcome and primary biological outcome for the Trial.

Defining the primary behavioral endpoint

Six principal challenges were encountered while developing a valid and reliable primary behavioral outcome: (i) the validity and reliability of self-report data; (ii) non-response to sensitive questions; (iii) interviewing strategy; (iv) recall window; (v) timing of the outcome assessment; and (vi) defining the primary behavioral outcome.

Validity and reliability of self-report data

The immediate goal of the Trial is to reduce sexual behaviors that potentially increase the risk of acquiring STD, including HIV. Of necessity, sexual behaviors must be self-reported. The validity of sexual behavior reports by individuals or aggregates of individuals depends on the willingness of the participants to respond to questions about their behavior honestly and their ability to recall specific behaviors accurately. The validity of the recorded information is influenced by many variables, including the memory of the participant, the context in which the information was elicited, the cultural mores of the social group to which the participant belongs, the level of social desirability bias, sex, the manner in which the information was obtained, and the participant's confidence in the researchers and their staff [4]. Using a standard protocol and a common assessment modality to interview individuals of similar age across different cultures may favorably influence the comparability of self-reports across countries. The comparability of cross-cultural reporting, however, especially among countries in diverse areas of the world, may be greatly influenced by specific cultural norms. For example, in societies in which it is considered taboo to discuss sex, convincing people to reveal their sexual activities accurately is difficult. Furthermore, any disclosure of

sexual promiscuity or unacceptable practices may be unlikely unless the participant is convinced that the interview will be anonymous.

Selection bias and non-response to sensitive questions

The willingness of individuals to participate in behavioral studies or answer sensitive questions about private behaviors is influenced by the cultural setting in which they live. For example, in China it is considered impolite to discuss one's sexual behavior, and any individual who does so would be considered 'odd' and likely be shunned by his or her peers. In other societies, such as India, it is acceptable for men to discuss their sexual activities, but it is unacceptable for women, as these discussions signal that a woman is immoral.

For these reasons, missing data may also be a problem for surveys in which sensitive information is elicited. Individuals who are willing to participate or answer specific sensitive questions may not be representative of the general population, and may include individuals who have low-risk behaviors and, thus, have less fear of disclosure. On the other hand, 'macho' individuals may be quite eager to participate and boast about their sexual prowess. In the Trial, the wording of the consent form, the contextual statements preceding the sensitive questions, and the sensitive questions themselves were field-tested to ensure as high a response rate as possible.

Interviewing strategy

To increase the chance of collecting valid information about sexual behaviors, it is necessary to gain the participant's confidence. Different interviewing methods have been used and compared to determine which strategies may best elicit sensitive information: (i) self-administration of a questionnaire; (ii) computer-assisted personal interviews (CAPI); or (iii) audio computer-assisted self-interviews (ACASI). No data had, however, been collected in these five countries on the effectiveness of these interviewing techniques. Therefore, before implementing the Trial, two small pilot studies were conducted in each country to assess the feasibility and reliability of using CAPI and ACASI to administer the behavioral assessment questionnaire (see 'The feasibility of audio computer-assisted self-interviewing in international settings' in this issue [18]). These strategies were considered because participants do not have to be literate to respond to questions. In China, and to a lesser extent in India, self-reports of behavior appeared somewhat less reliable using ACASI compared with CAPI. Therefore, CAPI was selected for the Trial in all five countries because the assessment administration must be standardized across sites.

Recall window

Participants who are able and willing to answer all questions honestly may nonetheless have difficulty accurately remembering activities that occurred over long, retrospective time frames. Most studies of sexual behavior attempt to use some strategy for estimating the frequency of sexual activity during a defined period in the past, introducing the possibility of recall bias. The extent of recall bias is influenced by the cognitive competence of the participant, which has been shown to be related to age, level of education, the consumption of alcohol and other drugs, and methods of eliciting the information [4]. In addition, recall ability may be affected by the frequency of the behavior and whether the activity was pleasant, unpleasant or neutral [19–21].

In a study comparing diary entries with sexual behaviors reported from interviews at one, 2, and 3 months after diary completion; recall bias was greater at 3 months compared with one month for one frequently occurring behavior, but not for other outcomes assessed [19]. Whereas long retrospective intervals (e.g. 12 months) may lead to unreliable data, low frequency behaviors may not occur during short recall windows. Three months has been suggested as a relatively reliable recall period [22]. Therefore, to limit recall bias in the Trial, participants are asked about behavior during the 3 months and 6 months before each assessment.

Timing of outcome assessment

Community-based interventions have been used across a range of health areas, including cardiovascular disease, HIV, and cancer, with varying results. Whereas many programmes aimed at HIV prevention appear to be effective, generally only modest effects have been seen for many studies in other programme areas. Reasons for the low impact achieved by many studies include methodological issues, such as low statistical power, the influence of naturally occurring changes in societal attitudes, and behavior that affects both control and intervention communities, smaller than anticipated effect sizes, an inadequate theoretical basis, and limitations of the intervention (e.g. insufficient dissemination throughout the community and length of the intervention) [23]. An ideal period in which change may be achieved and impact detected is not yet apparent. For example, researchers involved with the Community Intervention Trial for Smoking Cessation (COMMIT) speculated that 4 years was not enough to influence heavy smokers [24,25]. In contrast, several community-based HIV prevention studies have shown significant intervention effects on one or more outcomes after intervention periods lasting from one to 3 years [14,26,27].

On the basis of the data from these large trials and the performance of social diffusion interventions in the United States, data were collected at baseline and 12 and 24 months later to assess the behavioral and biological outcomes for the Trial.

Defining the primary behavioral outcome

During the ethnographic study conducted before beginning the Trial, qualitative interviewing supported the observation, also well supported in the literature, that sexual behavior in men and women varies substantially based on the perception of the type of partnership [28–30] [see 'Cross-site ethnographic findings that contributed to the design of the community popular opinion leader intervention in a five-country intervention study' in this issue [31] for a discussion of these differences among the Trial study populations]. Preliminary data indicated that the repertoire of behaviors was different with casual partners than with regular partners or spouses. In most of the Trial countries, women, especially married women, reported significantly fewer sex partners than men. On the basis of these ethnographic observations, we anticipated that we would not successfully change the sexual behaviors of partners within a marital relationship that had been well established for many years.

The primary behavioral outcome was established as a change in unprotected sex acts with non-spousal, non-live-in partners at 24 months. A change in condom use with non-spousal, non-live-in partners is, however, primarily a change in husbands' behaviors, which has the potential to be reflected in the rates of STD among their wives. Whereas a secondary behavioral outcome will be the change in unprotected acts with non-spousal, non-live-in partners at 12 months, the primary behavioral outcome was defined at 24 months because of concerns that the intervention may take longer than one year to disseminate throughout the community and produce a community-wide effect.

Primary biological endpoint

Three challenges emerged while developing a valid and reliable primary biological outcome: (i) variation in STD across sites; (ii) defining incidence; and (iii) ensuring the successful treatment of STD at baseline.

Variation in sexually transmitted diseases across sites

When evaluating prevention programmes across multiple countries, it is necessary to recognize the variation in culture and demography within and between Trial populations and sites [32]. These variations can result in different risk populations and different prevalent STDs. For example, herpes simplex virus-2 (HSV-2) is the most common STD in Zimbabwe [33], and is commonly related to co-infection with trichomonas, HIV, and bacterial STD in young adults [34,35]. In China, chlamydia is one of the most prevalent STDs, and the rates of STDs are highest among middle and upper-class migrant Chinese in eastern coastal cities, such as the owners and staff in the Fuzhou markets [36–38].

Preliminary epidemiological surveys conducted before beginning the Trial [39] showed significant variations in the prevalence of specific STD across the subpopulations in the countries included (Table 1; see 'Sexually transmitted disease and HIV prevalence and risk factors in concentrated and generalized HIV epidemic setting' [40] in this issue for further discussion). Using data from the first epidemiological study in China, Russia, and Zimbabwe, and from the second epidemiological study in India and Peru, the percentage of HIV-positive participants in each venue in Zimbabwe was on average 26%, whereas the average was below 3% in each of the other four countries. HSV-2 was the most common STD in Zimbabwe, Peru, and India (average percentage positive 45, 30, and 24%, respectively), but prevalence was below 10% in most venues in China and Russia. For gonorrhea, chlamydia, and syphilis, the average percentage of positive participants in the venues was below 10% in all countries. Among women, the prevalence of trichomonas was highest in India, where the average percentage positive across venues was 22%.

A challenge was thus defining a primary biological outcome to which significant disease levels from all country sites could contribute. Although differences were found in the prevalence of specific STD across the

Table 1. Bacterial and viral sexually transmitted diseases reported by participants in China, India, Peru, Russia, and Zimbabwe.

	China			India			Peru			Russia			Zimbabwe		
	N	Mean[a]	Range[a]	*N*	Mean	Range	*N*	Mean	Range	*N*	Mean	Range	*N*	Mean	Range
HIV	1529	0.0	0–0	1355	2.9	0–17	1194	1.5	0–11	991	0.6	0–6	1594	25.6	10–42
HSV-2	1529	8.6	2–16	1356	23.7	12–37	1194	29.9	14–54	941	6.2	0–18	1574	44.7	24–66
Syphilis	1520	1.6	0–4	1354	5.8	0–13	1192	5.5	0–20	987	0.6	0–4	1596	1.8	0–8
Chlamydia	1478	8.4	2–20	1348	1.0	0–7	1199	6.8	0–17	985	8.0	0–32	1284	2.1	0–7
Gonorrhea	1485	0.9	0–4	1347	0.4	0–2	1199	0.8	0–6	988	1.0	0–2	1279	1.2	0–8
Trichomonas	671	6.9	0–24	153	22.2	0–50	109	6.4	0–67	399	0.5	0–5	905	15.4	3–42
Venues	30			24			26			20			32		

HSV-2, Herpes simplex virus type 2.

[a]Mean and range across venues for percentage of participants with each sexually transmitted disease from a sample of approximately 50 participants in each venue. (See 'Selection of populations represented in the NIMH Collaborative HIV/STD Prevention Trial' [40] in this issue for a description of venues in each country.) Data are from the first epidemiological study in China, Russia, and Zimbabwe, and from the second epidemiological study in India and Peru.

countries, each of the sexually transmitted infections assessed are acquired through a common behavior, unprotected sex with an infected partner. The investigators thus considered the acquisition of any new sexually transmitted disease, or the re-occurrence of any effectively treated STD, as a biological outcome indicative of unprotected sex with an infected partner. Because the goal of the Trial is to assess the affects of the intervention on sexual behaviors, partner choice and condom use, this composite STD endpoint, comparing the rates of acquisition of new STDs between intervention and comparison subjects, provides a single biological outcome that can be used across sites with variable incidence of different STDs.

Defining incidence

In designing measures of disease incidence, the pathogenesis, treatment and transmission dynamics of each STD needs to be considered as well as the characteristics of the biological assays that define a new infection. Bacterial and protozoal STDs, such as gonorrhea, chlamydia, and trichomonas, are curable. Effective treatment reduces the period during which individuals with these diseases are infectious, while increasing the period during which they are susceptible to new infections. Syphilis can also be treated; and well-defined serological tests differentiate long-standing infection, treated or untreated, from newly acquired infection. Individuals infected with syphilis, unlike individuals infected with gonorrhea or chlamydia, may enter into a latent stage of disease 1–4 years after infection with relative immunity to re-infection and a lower risk of infecting a sex partner. Ignoring the differences in the biology of infections such as syphilis compared with diseases such as gonorrhea can lead to inaccurate assessments of the impact of a behavioral intervention [41].

In contrast, viral STDs such as HSV-2 or HIV cannot be cured, and although treatment may reduce infectiousness, both treated and untreated individuals may remain infectious and do not acquire a new infection with the same pathogen. Determining the incidence of viral infections is less problematical than for bacterial infections because the possibility of cure and re-infection is moot.

The combined biological STD endpoint provides a reliable method, adapted to the variations in disease prevalence in different countries, to identify the status (susceptible or infected) of each individual entering the Trial for each of six sexually transmitted infections (chlamydia, gonorrhea, HIV, HSV-2, syphilis and trichomonas). By repeating all of the biological measures at each study visit over 2 years, an incidence of STD may be calculated for each individual, community and country, as described in more detail below. The aggregate incidence of new STDs (percentage of new infections over time), and the comparison between intervention and comparison venues is presented as a 'combined' biological endpoint, less susceptible to reporting bias, and providing an accurate indication of a complex set of behaviors and interactions. Assessment of the prevalence of each infection in the target populations during the epidemiological studies before the start of the Trial contributed to estimates of the power of the Trial to measure statistically significant differences between intervention and comparison arms with regard to incident STD.

As described in the 'Methodological overview of a five-country community level HIV/sexually transmitted disease prevention trial' in this issue [17], an individual will be classified as a new case if they tested positive at either the 12 or 24-month follow-up for chlamydia, gonorrhea, syphilis (if negative at baseline), trichomonas (if female), HIV (if not positive for HIV at baseline), or HSV-2 (if not positive for HSV-2 at baseline). In addition, a person who tested positive for syphilis at both baseline and 24 months will be considered a new case if it can be demonstrated that the syphilis at baseline was treated, the follow-up test at 12 months was negative, and the 24-month result was positive. An individual who was positive for syphilis at baseline, 12 and 24-month follow up and who had a fourfold rapid plasma reagin titer decline from baseline to 12-month follow-up and a fourfold rapid plasma reagin titer increase from 12 to 24-month follow-up will be considered a new case. Table 2 characterizes the definition of a new case of syphilis for specific scenarios.

An individual will be classified as negative for the composite outcome if at least two-thirds of the tests used in the individual's assessment are non-missing and all provide definitive negative results (i.e. negative or indeterminate). If there are no new positive tests and more than a third of the individual's test results are missing, the composite variable will be set to missing.

Ensuring successful treatment of sexually transmitted diseases at baseline

Unique cultural and political factors impact healthcare-seeking behavior, the availability and acceptance of treatment, and the success of efforts to prevent re-infection and further transmission, including partner treatment or partner notification. For example, all STDs are reportable in Russia, even those discovered during conduct of a research project. Therefore, to measure a biological STD outcome in Russia it is necessary to assess the STD while concealing the identity of participants so that the research does not stigmatize an individual. The participant's disease, but not name, may be reported as required. In contrast, people from India and China frequently seek treatment of STD at pharmacies that specialize in treatments using local herbal remedies. To assess STD reliably in India and China over time, it is necessary to ensure that participants receive treatment from healthcare providers who will treat the STD with recommended medicines. Treatment protocols were established for the Trial to increase the likelihood that

Table 2. Definition of a new syphilis case under various scenarios.

Syphilis results			
Baseline	1st Follow-up	2nd Follow-up	New case since baseline
Negative (−)	Positive (+)	Positive (+)	Yes
Negative (−)	Positive (+)	Negative (−)	Yes
Negative (−)	Negative (−)	Positive (+)	Yes
Negative (−)	Negative (−)	Negative (−)	No
Positive (+)	Positive (+)	Positive (+)	Yes if fourfold decline from baseline to 12-month follow-up and fourfold rise from the 12 to the 24-month follow-up, else, No
Positive (+)	Positive (+)	Negative (−)	No
Positive (+)	Negative (−)	Positive (+)	Yes
Positive (+)	Negative (−)	Negative (−)	No

appropriate counseling and treatment were received by participants.

Treatment data are captured on a participant summary report and a symptom questionnaire. Both documents are completed at baseline and at each follow-up assessment. The Trial protocol requires that a participant summary report be completed for each participant who provides biospecimens at each visit. This report captures specific information such as whether laboratory test results were reported to the participant, the date those results were reported, and an explanation of why results were not reported if this is the case. More importantly, for participants who receive positive laboratory test results for any of the six STDs, the participant summary report indicates whether the participant received treatment, was referred for treatment to a medical facility, received education, or no action was taken.

The symptom questionnaire is administered to all participants who provide biological specimens at each assessment. The symptom questionnaire captures information regarding whether treatment for an STD was received in the past 12 months (China and Russia only), and whether the participant has experienced genital discharge, sharp/burning pain during urination, or genital sores in the past week. If the participant's responses to the symptom questionnaire indicate that he or she has STD symptoms, free treatment or a referral for free treatment is given. Sites then record whether participants who presented with symptoms were referred for treatment or treated for an STD by project staff. The medications used for treatment are also captured on this form. During this interaction, counselors stress the importance of the participant returning for the results of the laboratory tests, even if treatment is received during the assessment. Counselors explain that many people who have an STD do not have symptoms or the symptoms reported do not identify the correct type of STD. Only by returning for test results can the participant receive an accurate diagnosis and additional treatment or referral, if necessary.

Data collected from the participant summary report and the symptom questionnaire, in conjunction with laboratory results, will determine if participants with positive laboratory results received treatment or referral for the correct disease(s). Review of treatment received, endpoints, and laboratory testing will help differentiate incident and prevalent cases of bacterial and protozoal infections.

Discussion

Many researchers have strong biases about the utility of behavioral versus biological endpoints in assessing the impact of interventions on public health [42]. Identifying a set of outcome measures that can be used across countries and with different populations is a complex issue not easily resolved. A reduction in STD incidence is potentially a valid and reliable endpoint that can be applied across a variety of settings. Even STD endpoints are influenced by social and cultural factors, such as sexual norms, healthcare-seeking behavior, knowledge about modes of transmission and the correct use of condoms, and the availability of treatment, and the use of individual STD endpoints may only be feasible in high-prevalence and incidence settings.

Behavioral outcomes are the immediate target of behavioral interventions, and may be used in low-prevalence countries. Behavioral measures, however, do not always predict incident STD, and may be subject to a different set of cultural influences causing bias. In addition, behavioral measures may not enable policymakers and those evaluating interventions to estimate the direct impacts of the interventions on HIV/STD incidence. Essentially, without knowing how many infections can be prevented, knowing the number of individuals necessary to treat in order to avert a certain number of infections, or having the ability to compare the impact of a variety of prevention interventions, studies of interventions without biological outcomes are of limited merit to decision-makers.

In addition, mathematical modeling of the epidemic can prove a powerful tool to overcome this problem by assessing both where the treatment and prevention funds should be spent and what kind of prevention programme designs will potentially avert the most cases of new infections. Multiple mathematical models of the dynamics of the HIV epidemic can be used in planning and designing HIV/STD prevention programmes for emerging epidemics in developing countries. For example, the NIMH/ASIST (AIDS Strategic Interventions Simulation Tool) was developed specifically to assess the impact of different behavioral intervention strategies on the incidence of HIV and STD [43].

In conclusion, a challenge in implementing the Trial was to develop behavioral and biological primary outcome measures that can be used to evaluate the effectiveness of the C-POL intervention implemented in diverse populations and settings in five countries. Researchers wanted to evaluate both the immediate outcome of the intervention on sexual behavior and provide a measure of impact that is biologically relevant and relatively resistant to self-disclosure bias. Accordingly, both behavioral and biological endpoints were defined: (i) a change in self-reported unprotected sex acts with non-spousal, non-live-in partners during the 3 months before the 24-month assessment measured as a change from baseline; and (ii) the incidence of STD defined as any new case of chlamydia, gonorrhea, syphilis, trichomonas (women), HSV-2, or HIV at 12 or 24 months. To evaluate the C-POL intervention, a behavioral outcome was thus developed that accounts for the relationship between sexual behavior and the type of partner by targeting the partnerships in which change is thought to be most likely, reducing the potential for recall bias by limiting the recall period to 3 months, and allowing 2 years for the effects of the intervention to be measured. Similarly, a combined biological endpoint was developed that can accommodate the variation in prevalence of individual STD across countries.

Because these measures provide different information, both behavioral and biological endpoints should be considered for the evaluation of a wide variety of intervention trials, including vaccine, microbicide, and peer-education. Although implementing the assessment of both behavioral and biological outcomes in this Trial is more complex, it permits a more complete assessment of the effects of the intervention, which provides clues about how to scale up the intervention effectively so that it can have the intended public health impact.

References

1. National Institutes of Health. **Consensus development conference statement, February 11–13, 1997.** *AIDS 2000*; **14 (Suppl. 2)**: 85–96.
2. Boerma JT, Weir SS. **Integrating demographic and epidemiological approaches to research on HIV/AIDS: the proximate-determinants framework.** *J Infect Dis* 2005; **191 (Suppl. 1)**: S61–S67.
3. Fishbein M, Pequegnat W. **Evaluating AIDS prevention interventions using behavioral and biological outcome measures.** *Sex Transm Dis* 2000; **27**:101–110.
4. Pequegnat W, Fishbein M, Celentano D, Ehrhardt A, Garnett G, Holtgrave D, *et al*. **NIMH/APPC workgroup on behavioral and biological outcomes in HIV/STD prevention studies: a position statement.** *Sex Transm Dis* 2000; **27**:127–132.
5. Anderson RM, May RM. *Infectious diseases of humans: dynamics and control*. Oxford, UK: Oxford University Press; 1991
6. Anderson R. **Transmission dynamics of sexually transmitted infections.** In: *Sexually transmitted diseases,* 3rd ed. Edited by Holmes KK, Sparling PF, Mardh P-A, *et al*. New York: McGraw-Hill; 1999. pp. 25–37.
7. Royce RA, Sena A, Cates W Jr, Cohen MS. **Sexual transmission of HIV.** *N Engl J Med* 1997; **336**:1072–1078.
8. Grassly NC, Garnett GP, Schwartlander B, Gregson S, Anderson RM. **The effectiveness of HIV prevention and the epidemiological context.** *Bull WHO* 2001; **79**:1121–1132.
9. Aral SO, Holmes KK. **Social and behavioral determinants of the epidemiology of STDs: industrialized and developing countries.** In: *Sexually transmitted diseases,* 3rd ed. Edited by Holmes KK, Sparling PF, Mardh P-A, Lemon SM, Stamm WE, Piot P, *et al*. New York: McGraw-Hill; 1999. pp. 39–76.
10. Aral SO, Peterman TA. **A stratified approach to untangling the behavioral/biomedical outcomes conundrum.** *Sex Transm Dis* 2002; **29**:530–532.
11. Kelly JA. **Popular opinion leaders and HIV prevention peer education: resolving discrepant findings, and implications for the development of effective community programmes.** *AIDS Care* 2004; **16**:139–150.
12. Kelly JA, St Lawrence JS, Diaz YE, Stevenson LY, Hauth AC, Brasfield TL, *et al*. **HIV risk behavior reduction following intervention with key opinion leaders of a population: an experimental analysis.** *Am J Public Health* 1991; **81**:168–171.
13. Kelly JA, St Lawrence JS, Stevenson LY, Hauth AC, Kalichman SC, Diaz YE, *et al*. **Community AIDS/HIV risk reduction: the effects of endorsements by popular people in three cities.** *Am J Public Health* 1992; **82**:1483–1489.
14. Kelly JA, Murphy DA, Sikkema KJ, McAuliffe TL, Roffman RA, Solomon LJ, *et al*. **Randomized, controlled, community-level HIV prevention intervention for sexual risk behaviour among homosexual men in U.S. cities.** *Lancet* 1997; **350**:1500–1505.
15. Sikkema KJ, Kelly JA, Winett RA, Solomon LJ, Cargill VA, Roffman RA, *et al*. **Outcomes of a randomized community-level HIV prevention intervention for women living in 18 low-income housing developments.** *Am J Public Health* 2000; **90**:57–63.
16. Sikkema KJ, Anderson ES, Kelly JA, Winett RA, Gore-Felton C, Roffman RA, *et al*. **Outcomes of a randomized, controlled community-level HIV prevention intervention for adolescents in low-income housing developments.** *AIDS* 2005; **19**:1509–1516.
17. 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–S18.
18. NIMH Collaborative HIV/STD Prevention Trial Group. **The feasibility of audio computer-assisted self-interviewing in international settings.** *AIDS* 2007; **21 (Suppl. 2)**: S49–S58.
19. Graham CA, Catania JA, Brand R, Duong T, Canchola JA. **Recalling sexual behavior: a methodological analysis of memory recall bias via interview using the diary as the gold standard.** *J Sex Res* 2003; **40**:325–332.
20. Catania JA, Gibson DR, Chitwood DD, Coates TJ. **Methodological problems in AIDS behavioral research: influences on measurement error and participation bias in studies of sexual behavior.** *Psychol Bull* 1990; **108**:339–362.
21. Skowronski JJ, Betz AL, Thompson CP, Shannon L. **Social memory in everyday life: recall of self-events and other-events.** *J Personal Soc Psychol* 1991; **60**:831–843.
22. Kauth MR, St Lawrence JS, Kelly JA. **Reliability of retrospective assessments of sexual HIV risk behavior: a comparison of biweekly, three-month, and twelve-month self-reports.** *AIDS Educ Prev* 1991; **3**:207–214.

23. Merzel C, D'Afflitti J. **Reconsidering community-based health promotion: promise, performance, and potential.** *Am J Public Health* 2003; **93**:557–574.
24. COMMIT Research Group. **Community Intervention Trial for Smoking Cessation (COMMIT): I. Cohort results from a four-year community intervention.** *Am J Public Health* 1995; **85**: 183–192.
25. COMMIT Research Group. **Community Intervention Trial for Smoking Cessation (COMMIT): II. Changes in adult cigarette smoking prevalence.** *Am J Public Health* 1995; **85**:193–200.
26. **Community-level HIV intervention in 5 cities: final outcome data from the CDC AIDS Demonstration Projects.** *Am J Public Health* 1999; **89**:336–345.
27. Lauby JL, Smith PJ, Stark M, Person B, Adams J. **A community-level HIV prevention intervention for inner-city women: results of the Women and Infants Demonstration Projects.** *Am J Public Health* 2000; **90**:216–222.
28. Morris M, Pramualratana A, Podhisita C, Wawer MJ. **The relational determinants of condom use with commercial sex partners in Thailand.** *AIDS* 1995; **9**:507–515.
29. Norris AE, Ford K, Shyr Y, Schork MA. **Heterosexual experiences and partnerships of urban, low-income African-American and Hispanic youth.** *J Acquir Immune Defic Syndr* 1996; **11**:288–300.
30. Macaluso M, Demand MJ, Artz LM, Hook EW III. **Partner type and condom use.** *AIDS* 2000; **14**:537–546.
31. NIMH Collaborative HIV/STD Prevention Trial Group. **Formative study conducted in five countries to adapt the community popular opinion leader intervention.** *AIDS* 2007; **21 (Suppl. 2)**: S91–S98.
32. Manhart LE, Holmes KK. **Randomized controlled trials of individual-level, population-level, and multilevel interventions for preventing sexually transmitted infections: what has worked?** *J Infect Dis* 2005; **191 (Suppl. 1)**:7–24.
33. Decosas J, Padian N. **The profile and context of the epidemics of sexually transmitted infections including HIV in Zimbabwe.** *Sex Transm Infect* 2002; **78 (Suppl. 1)**:40–46.
34. Decosas J. **Labour migration and HIV epidemics in Africa.** *AIDS Anal Africa* 1998; **8**:6–7.
35. Hope KR Sr. **Mobile workers and HIV/AIDS in Botswana.** *AIDS Anal Africa* 2000; **10**:6–7.
36. Detels R, Wu Z, Rotheram MJ, Li L, Guan J, Yin Y, *et al*. **Sexually transmitted disease prevalence and characteristics of market vendors in eastern China.** *Sex Trans Dis* 2003; **30**:803–808.
37. He N, Detels R, Zhu J, Jiang Q, Chen Z, Fang Y, *et al*. **Characteristics and sexually transmitted diseases of male rural migrants in a metropolitan area of eastern China.** *Sex Transm Dis* 2005; **32**:286–292.
38. Parish WL, Laumann EO, Cohen MS, Pan S, Zheng H, Hoffman I, *et al*. **Population-based study of chlamydial infection in China: a hidden epidemic.** *JAMA* 2003; **289**:1265–1273.
39. NIMH Collaborative HIV/STD Prevention Trial Group. **Sexually transmitted disease and HIV prevalence and risk factors in concentrated and generalized HIV epidemic settings.** *AIDS* 2007; **21 (Suppl. 2)**: S81–S90.
40. NIMH Collaborative HIV/STD Prevention Trial Group. **Selection of populations represented in the NIMH Collaborative HIV/STD Prevention Trial.** *AIDS* 2007; **21 (Suppl. 2)**: S19–S28.
41. Grassly NC, Fraser C, Garnett GP. **Host immunity and synchronized epidemics of syphilis across the United States.** *Nature* 2005; **433**:417–421.
42. Schacter J. **Biologic versus behavioral endpoints: the duet continues.** *Sex Transm Dis* 2000; **27**:456–457.
43. NIMH/ASIST Program Committee. *NIMH/ASIST (AIDS Strategic Interventions Simulation Tool). Manual for beginners and experts*. Bethesda, MD: National Institute of Mental Health; 2000.