Non-viral STIs

....

Design, measurement, and analytical considerations for testing hypotheses relative to condom effectiveness against non-viral STIs

R Crosby, R J DiClemente, D R Holtgrave, G M Wingood

....

Recently, the US Department of Health and Human Services (DHHS) issued a report on a workshop that synthesised evidence regarding the effectiveness of latex condoms for the prevention of sexually transmitted infections (STIs).[1] The report cited evidence that condoms are effective in preventing HIV transmission and female to male transmission of gonorrhoea, but stated that empirical evidence was insufficient to evaluate the degree of risk reduction provided by condoms with regard to chlamydia, syphilis, chancroid, trichomoniasis, genital herpes, and human papillomavirus. One important implication of the report is that there is a need for further research on condom effectiveness. As the report noted, "to definitively answer the remaining questions about condom effectiveness for preventing STD infections will require well designed and ethically sound clinical studies."[1]

Besides the research perspective, from an applied public health perspective, intensified efforts to test condom effectiveness are urgently needed. Firstly, the DHHS report may have eroded public confidence in an otherwise widely recommended method of STI prevention (see Centers for Disease Control and Prevention, 1996 for public recommendations[2]). In addition, if confidence in the effectiveness of condoms declines among health professionals and other policy makers, then their efforts to promote condom use also may wane. Consequently, people at risk of STI infection may be less likely to adopt or sustain condom use as a STI prevention strategy.

Given the in vitro evidence that intact latex condoms are virtually impermeable to even the smallest of STI pathogens,[3–5] the present lack of in vivo evidence supporting condom effectiveness against many STIs should not be counted as evidence that condoms are ineffective. Numerous, complex challenges are inherent in the design and analysis of in vivo tests of condom effectiveness.

In this editorial, we describe selected key issues that should be addressed and resolved in tests of hypotheses relevant to condom effectiveness for *non-viral* STI prevention. The issues we have selected are not meant to comprehensively represent all possible issues that should be considered in tests of condom effectiveness, but are those that we feel are central. These issues involve design, measurement, and data analytical procedures that are used to test hypotheses related to condom effectiveness (that is, the protective value—calculated from cohort studies and often expressed as an odds ratio—of condom use against acquisition of non-viral STIs).

DESIGN ISSUES

Serodiscordant couple studies have been useful in establishing the strong protective value of condoms against HIV infection.[6 7] A similar approach could not ethically be applied to investigations involving non-viral STIs because doing so would involve withholding medical treatment. However, it is possible to create prospective cohort study designs that ethically address condom effectiveness against non-viral STIs.

Suggested prospective design

In a well designed prospective cohort study of condom effectiveness, an initial assessment is needed to establish an infection free cohort of individuals. This assures investigators that all STIs detected at follow up are true incident infections. In an era of highly sensitive and specific tests for many common STIs, preferable outcome measures for tests of condom effectiveness are STIs that can be assessed by nucleic acid amplification tests. In particular, recent microbiological advances have developed nucleic acid amplification assays for two prevalent STIs (chlamydia and gonorrhoea) that could provide reliable outcome measures for tests of condom effectiveness. Therapeutic advances also contribute to increased methodological rigour, as many non-viral STIs can be effectively treated with single dose orally administered therapy (that is, treatment occurs at the initial assessment). Direct observation of this therapy provides increased assurance that the entire cohort begins the prospective phase of the study without active infections.

In subsequent (follow up) phases of the prospective study both the predictor variable (that is, retrospective reports of condom use) and the outcome measures (that is, incident STIs) should be assessed. Incident STIs should be assessed using identical assays and diagnostic evaluation criteria. This concurrent collection of predictor and outcome variables will not confound the design as long as the condom use data are collected before participants discover their STI test results.

Study designs should also be planned to provide at least 80% power to detect small to modest effect sizes. Planned sample sizes should conservatively allow for a retention rate of 70%. When studies of condom effectiveness are inadequately powered (as many are), an alternative approach to significance testing is to report the obtained effect size (that is, based on the difference between groups, calculate and report the effect size obtained for condom use). Studies of small samples may achieve modest to large effect sizes despite their lack of power to achieve significance as defined by traditional 95% confidence intervals (pooling these studies for meta-analysis may then be appropriate).

Temporal ordering

An unfortunate and potentially irresolvable limitation of this design is that it cannot discriminate between condom use as a behavioural response to perceived threat of STI *acquisition* (a self protective behaviour) and condom use as a means of preventing STI *transmission* (a partner protective behaviour). Given that biological testing is conducted at two time points, the retrospective recall period for condom use, by necessity, corresponds to the entire time between the points rather than between time one and infection. Although an argument could be made that this problem is overcome by daily or weekly testing, the behavioural effects of this repeated testing are highly likely to confound a fair test of the hypothesis that condoms are protective against non-viral STIs.

Because these study designs are not capable of determining what proportion of an individual's condom use followed acquisition of an STI, we suggest that only two-tailed tests of significance should be used in tests of hypotheses relative to condom effectiveness. Using a two-tailed test, rejecting the null hypothesis (that is, no association) would allow for two possibilities: (1) an *inverse* association suggesting that condom use prevented STI acquisition, or (2) a *direct* association suggesting that condom use was a behavioural response to suspected

STI acquisition. Specifically ruling out the latter possibility is necessary to establish the former.

MEASUREMENT ISSUES
Measurement of condom use frequency

How condom use is measured and the length of the recall period are two important factors that can have a profound effect on the strength of association between condom use and STI incidence. In condom research, two distinct measures of condom use (or non-use) have been employed: proportional (percentage times a condom was used) and absolute (the number of times a person reported unprotected sex). Recent publications support the value of using an absolute measure of condom use.[8–11]

A proportional measure of condom use is created by dividing the number of condom protected sex acts by the total number of sex acts. Unfortunately, this division eliminates variance across study participants relative to the frequency of intercourse during the selected recall period. Thus, proportional measures fail to capture variance related to abstinence from sex if they are used in isolation from information capturing frequency of intercourse. For example, consider person "A" who used condoms 90% of the time for 100 episodes of intercourse and person "B" who used condoms 50% of the time for 10 episodes of intercourse. Using a proportional measure would suggest that person B (50% use) is at high risk relative to person A (90% use) when in fact, person B has half as many potential exposures to STIs compared to person A. Thus, a proportional measure condom use may underestimate the true risk for STIs as it fails to capture the number of potential exposures. In contrast, subtracting the number of condom protected sex acts from the total number of sex acts) yields a variable that reflects only the number of sexual occasions that were not condom protected.

The second factor to consider is the length of the recall period (that is, the length of time between assessment points). An extensive empirical literature has addressed the relative merits of using different recall periods to assess condom use.[10 12–14] Yet, consensus on an optimal length for recall periods has not been achieved. Consequently previous studies have utilised a variety of period lengths ranging from the last sexual act to a retrospective period of 6 months.[10] We suggest that the length for recall periods should match the time elapsing between assessments in prospective studies of condom effectiveness. Thus, a primary design question (that is, how much time to allow between assessment periods) can be resolved by carefully considering the pros and cons of various recall periods and selecting a period that best suits the study population.

Several considerations are relevant to selecting the length of recall periods used in studies of condom effectiveness. For instance, longer recall periods are problematic because they increase the likelihood of inaccurate reporting. Conversely, shorter recall periods necessitate frequent screening which, in turn, increases participant response burden and the cost of the study. Shorter recall periods may also limit the available sample size in that some members of the cohort may not be sexually active during the shorter time frame, thus it would be inappropriate to include these study participants in analyses pertaining to the corresponding assessment period. Of course, tests of condom effectiveness could use a "proxy" recall period (for example, the last 30 days of a 6 month time frame are used as the recall period; responses are assumed to reflect the entire 6 months). While this approach may minimise inaccurate recall, it may not adequately reflect condom use between the two STI screenings (for example, baseline and 6 month follow up).

One additional measure of condom use, commonly employed, does not utilise a time period, per se, but rather simply asks participants whether they used a condom the last time they had sex. In this case, participants did or did not use a condom. Thus, condom use can only be quantified as either 0% (did not use at last intercourse) or 100% (did use at last intercourse). Although this measure yields optimal accuracy of recall and is attractive based on its simplicity, its utility is limited. Naturally, the "last act" measure captures only a single coital episode and this may poorly reflect participants' condom use across the entire period of follow up.

Measurement of condom use errors and problems

As is true with measures of contraceptive effectiveness (for example, hormonal contraception, natural family planning), measures of condom effectiveness can be conceptualised at two distinct levels: (1) effectiveness given perfect (flawless) use, and (2) effectiveness given typical use.[15] We suggest that in vivo tests of condom effectiveness should account for variance across condom users relative to imperfect use of latex condoms. Recent evidence suggests that this variance may be substantial.[11 16–21] For example, a recent study of 158 condom using college men found that 43% reported recently putting a condom on after starting sex and 15% recently reported taking a condom off before sex was over.[22] The study also found that 30% placed the condom upside down on the penis and had flipped the condom over before starting sex (thereby potentially transferring infected seminal fluid to the exterior of the condom). The authors of this study created an assessment instrument that quantified 24 errors (including the three described here) that could compromise condom effectiveness. In addition, the authors identified and assessed four likely problems or outcomes of these errors, including slippage and breakage (reported by 35%). Thus, given that wide range of condom use errors and problems that can compromise condom effectiveness, any test of condom effectiveness that does not include comprehensive measures of errors and problems could severely underestimate perfect use and misrepresent typical use. Stated differently, tests of condom effectiveness must first rule out user error before conclusions can be made about product error.

Validity

Measurement of condom use necessitates an assumption that study participants accurately recall and report their frequency of sex and use of condoms. Clearly, this assumption poses problems. For example, Zenilman and colleagues noted that over-reporting of condom use may have explained an observed lack of association between condom protected sex and STI incidence among a sample of adolescents.[23] Subsequent research efforts have tested strategies designed to elicit more accurate recall and improved disclosure.

Recalling frequency of sex and condom use may be facilitated by methods that involve daily recording such as keeping diaries.[24] A useful practice is to design questions prompting people to consider sex and condom use with steady partners as well as with non-steady partners. Researcher(s) should also clearly specify (in the data collection instrument) a clear definition of sex (for example, "sex means putting the penis in the vagina") and condom use (for example, "condom use means using a latex condom from start to finish of sex without having it break or fall off"). Keeping the number of questions to a minimum may also prevent undue respondent burden thereby promoting improved effort (which, in turn, may promote accurate recall). Given that only one recall period is used (that is, that corresponding to the time between assessment intervals), condom use can be assessed with a relatively short list of items. These items should include frequency of sex and condom use (considering all sex partners), whether any "fatal" condom use errors occurred (for example, breakage, slippage during sex or withdrawal, partial use, pre-sex contamination of condoms with semen), and whether latex condoms were used.

Honest disclosure may also be facilitated by several strategies. Strategies may increase confidentiality—for example, by use of audio computer assisted

self interviewing.[25] Disclosure may also be enhanced by developing questionnaires based on preliminary interactions with members of the target audience.[26] However, it is important to note that empirically establishing the utility of these strategies (as well as those designed to promote recall) is problematic owing to a lack of "gold standards" for comparison purposes. An area of needed further research is developing and refining assays that can be used to confirm study participants' reports of having unprotected sex.

ANALYTICAL ISSUES
Skewed distributions and effect size

A common, and yet to be resolved, problem in condom research is that distributions of scores reflecting participants' condom use may grossly fail to meet necessary assumptions for parametric analyses. In highly skewed distributions (that is, distribution with a long tail, often created by extreme scores), standard deviations tend to be quite large; thus effect size is reduced. Absolute measures of condom use are vulnerable to inflated standard deviations created by extreme scores. Conversely, proportional measures are not prone to extreme scores because the range of possible values has been mathematically constrained (that is, measurement is converted to a percentage between 0 and 100). Thus, the more conservative approach to testing hypotheses of condom effectiveness is to evaluate absolute, as opposed to proportional, measures of condom use.

Dichotomising condom use measures

A second, and related, analytical issue is whether and how to dichotomise skewed distributions of absolute or proportional measures of condom use. Parametric analyses may misrepresent data because distributions representing proportional or absolute measures of condom use tend to be highly skewed, curvilinear, or even bimodal. Although logarithmic transformations have been applied to condom use measures, one unavoidable problem with this treatment of data is that interpretation of the transformed values becomes problematic. Alternatively, dichotomising non-normally distributed condom use measures provides improved representation of the data and the findings can easily be expressed in common epidemiological terms (for example, odds ratios).

Deciding how to dichotomise non-normally distributed condom use measures is, however, problematic. A frequently employed procedure has been to dichotomise these distributions using "never" versus "some/always" or "always" versus "some/never." Note that the

options only differ by which end of the dichotomy contains the "sometimes" (that is, scores representing 1% to 99% condom use). In study designs that test hypotheses relative to condom effectiveness, the primary research questions can be thought of as "How much does consistent (that is, 100%) condom use reduce the odds of STI acquisition?" Thus, the logical dichotomy is to compare "always" users to "some or never" users (for proportional measures) and participants classified as "never" engaging in unprotected sex versus those engaging in "any" unprotected sex (for absolute measures). Unfortunately, both procedures fail to provide information relative to a second, and potentially equally important, research question: "Does some use of condoms provide partial protection against STI infection?"

Of note, decisions about how to dichotomise proportional measures of condom use also have implications for survey research. Arbitrarily grouping the "sometimes users" with either "always" or "never" users may lead to type 1 and type 2 errors.[10] Instead, researchers may want to consider performing analyses that compare "sometimes users" to "always users" and to "never users." This process can be used to establish an empirical basis for creating the dichotomy—for example, if "sometimes users" do not differ from "never users" with respect to the variables under investigation, then combining the two groups and comparing them to "always" users is well justified.

Investigating dose-response relations

Using data from modelling exercises, Pinkerton and Abramson suggested that occasional condom use can be partially protective against HIV infection.[27] To the best of our knowledge, similar modelling exercises relevant to non-viral STIs have not been published. Whether by modelling or by observational studies, empirical evidence addressing a potential dose-response relation between condom use and non-viral STI prevention could prove quite interesting. Such research would begin to examine the relation of condom use to STI among "sometimes" users (that is, people with scores representing 1% to 99% condom use). For example, proportional measures of condom use could be divided into deciles and the percentage of people testing positive for STI in each of these deciles could be calculated.

One inherent limitation of testing for dose-response relations among "sometimes" users of condoms is that people may engage in situational condom use. In other words, people may use condoms when they have sex with partners whom they perceive to be potential transmitters

of an STI. Empirical support for this phenomenon can be found in an analysis of data obtained from Project RESPECT; the researchers concluded that, "people tend to have safe sex with risky partners and risky sex with safe partners."[28]

Multivariate analyses

A second inherent limitation of testing hypotheses relative to condom effectiveness for less than 100% proportional use or greater than zero acts of unprotected sex is that partner related variables must be accounted for in the analyses. For example, if condom use is situational (by partner type) then assessments including diverse partner related measures are needed (for example, number of sex partners, STI risk behaviour of sex partners). Alternatively, in studies assessing the relation of STI incidence to 100% proportional condom use or "never" having unprotected sex, assessment of partner related variables would not be necessary (as, for example, 100% use implies STI protection regardless of partner related measures). Thus, while bivariate analyses can be used to test hypotheses relevant to consistent condom use, multivariate analyses are necessary for testing hypotheses relevant to the effects of less than consistent condom use (or greater than zero acts of unprotected sex) on STI incidence.

Given that data can be quite expensive to collect and that a substantial portion of research participants in any study will report less than consistent condom use, studies designed to assess the effectiveness of condoms clearly should measure the entire spectrum of partner related variables that could also account for variance in STI incidence. This analytical issue is also a measurement issue. For example, the quality of any one research participant's condom use may vary from one partner to another. Measurement of condom use errors and problems must be *partner specific*, particularly if study participants are largely reliant on their sex partners for correct application and use of condoms as is often true with women in relationships with men.

CONCLUSIONS

We have described many considerations necessary for a fair test of hypotheses relative to condom effectiveness. Indeed, tests of condom effectiveness will require a level of rigour that incorporates appropriate design, measurement, and analytical steps. Although this level of rigour will be resource intensive, the potential benefits of creating a fair test of hypotheses relevant to condom effectiveness are warranted (for example, restored public and provider confidence in condom effectiveness for STI prevention). Conversely, given that consistent and correct use of condoms has been such a vital

strategy applied to STI prevention (with some evidence suggesting efficacy), the cost of *not* providing fair tests of these hypotheses could be prohibitively high.

. . . .

Authors' affiliations
R Crosby, R J DiClemente, D R Holtgrave, G M Wingood, Rollins School of Public Health, Department of Behavioral Sciences and Health Education, and Emory Center for AIDS Research, Atlanta, Georgia, USA
R J DiClemente, Emory University School of Medicine, Department of Pediatrics, Division of Infectious Diseases, Epidemiology, and Immunology and Department of Medicine (Infectious Diseases), Atlanta, Georgia, USA

Correspondence to: Richard Crosby, PhD, Rollins School of Public Health of Emory University, Department of Behavioral Sciences and Health Education, 1518 Clifton Road, NE, Room 542, Atlanta, GA 30322, USA; rcrosby@sph.emory.edu

REFERENCES

1 **United Sates Department of Health and Human Services**. Workshop summary: scientific evidence on condom effectiveness for sexually transmitted disease (STD) prevention. Available online at http://www.niaid.nih.gov/dmid/stds/condomreport.pdf. Accessed 5 July 2001.
2 **Centers for Disease Control and Prevention**. *Facts about condom and their use in preventing HIV infections and other STDs.* Atlanta, GA: United Sates Department of Health and Human Services, 1996.
3 **Lytle CD**. Lack of latex porosity: a review of virus barrier tests. *J Rubb Res* 1999;**2**:23–39.
4 **Lytle CD**. An in vitro evaluation of condoms as barriers to a small virus. *Sex Transm Dis* 1997;**24**:161–4.
5 **Lytle CD**. A sensitive method for evaluating condoms as virus barriers. *J AOAC Inf* 1997;**80**:319–24.
6 **De Vincenzi I**. A longitudinal study of human immunodeficiency virus transmission by heterosexual partners. *N Engl J Med* 1994;**331**:341–6.
7 **Saracco A**, Musicco M, Nicolosi A, et al. Man-to-woman sexual transmission of HIV: Longitudinal study of 343 steady partners of infected men. *J AIDS* 1993;**6**:497–502.
8 **O'Leary AO**, DiClemente RJ, Aral SO. Reflections on the design and reporting of STD/HIV behavioral intervention research. *AIDS Educ Prev* 1997;**9(S-A)**:1–14.
9 **Crosby RA**, DiClemente RJ, Wingood GM, et al. Correlates of unprotected vaginal sex among African American female teens: the importance of relationship dynamics. *Arch Pediatr Adolesc Med* 2000;**154**:893–9.
10 **Crosby RA**. Condom use as a dependent variable: measurement issues relevant to HIV prevention programs. *AIDS Educ Prev* 1998;**10**:448–57.
11 **Fishbein M**, Peguegnat W. Evaluating AIDS prevention interventions using behavioral and biological outcome measures. *Sex Transm Dis* 2000;**27**:101–10.
12 **Catania JA**, Gibson DR, Chitwood DD, et al. Methodological problems in AIDS behavioral research: influences on measurement error and participation bias in studies of sexual behavior. *Psychol Bull* 1990;**108**:339–62.
13 **Kelly JA**, Murphy DA, Washington CD, et al. The effects of HIV/AIDS intervention groups for high-risk women in urban clinics. *Am J Public Health* 1994;**84**:1918–22.
14 **McFarlane M**, St. Lawrence JS. Adolescents' recall of sexual behavior: consistency of self-report and effect of variations in recall duration. *J Adolesc Health* 1999;**25**:199–206.
15 **Cates W**. Contraception, contraceptive technology, and STDs. In: Holmes KK, Sparling PF, Mardh P, et al, eds. *Sexually transmitted diseases*. 3rd ed. New York: McGraw-Hill, 1999:1067–78.
16 **Crosby RA**, DiClemente RJ, Wingood GM, et al. Correct condom application among African American adolescent females: the relationship to perceived self-efficacy and the association to confirmed STDs. *J Adolesc Health* 2001;**29**:194–9.
17 **Langer LM**, Zimmerman RS, Cabral RJ. Perceived versus actual condom skills among clients at sexually transmitted disease clinics. *Pub Health Rep* 1994;**109**:683–7.
18 **Macaluso M**, Kelaghan J, Artz L, Austin H, et al. Mechanical failure of the latex condom in a cohort of women at high STD risk. *Sex Transm Dis* 1999;**26**:450–7.
19 **Mertz KJ**, Finelli L, Levine WC, et al. Gonorrhea in male adolescents and young adults in Newark, New Jersey: implications of risk factors and patient preferences for prevention strategies. *Sex Transm Dis* 2000;**27**:201–7.
20 **Warner L**, Clay-Warner J, Boles J, et al. Assessing condom use practices: Implications for evaluating method and user effectiveness. *Sex Transm Dis* 1998;**6**:273–7.
21 **Crosby RA**, Yarber WL. Perceived versus actual knowledge about correct condom use among US adolescents: results from a national study. *J Adolesc Health* 2001;**28**:415–20.
22 **Crosby RA**, Sanders S, Yarber WL, et al. Condom use errors and problems among college men. *Sex Transm Dis* (in press).
23 **Zenilman JM**, Weisman CS, Rompalo AM, et al. Condom use to prevent incident STDs: The validity of self-reported condom use. *Sex Transm Dis* 1995;**22**:15–20.
24 **Fortenberry JD**, Orr DP, Katz BP, et al. Sex under the influence: a diary self-report study of substance use and sexual behavior among adolescent women. *Sex Transm Dis* 1997;**24**:313–19.
25 **Turner CF**, Ku L, Rogers SM, et al. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science* 1998;**280**:867–71.
26 **Michaud PA**, Narring F, Ferron C. Alternative methods in the investigation of adolescents' sexual life. *J Adolesc Health* 1999;**25**:84–90.
27 **Pinkerton SD**, Abramson PR. Occasional condom use and HIV risk reduction. *J AIDS* 1996;**13**:456–60.
28 **Peterman TA**, Lin LS, Newman DR, et al. Does measured behavior reflect STD risk?: an analysis of data from a randomized controlled behavioral intervention study. *Sex Transm Dis* 2000;**27**:446–51.