AIDS and Behavior, Vol. 8, No. 1, March 2004 (© 2004)

Brief Communication

Condom Use as a Dependent Variable: A Brief Commentary About Classification of Inconsistent Users

Richard A. Crosby,[1,2,3,7] **William L. Yarber,**[3,4,5,6] **Stephanie A. Sanders,**[3,4,5,6] **and Cynthia A. Graham**[4,6]

Received Mar. 18, 2003; revised Sep. 4, 2003; accepted Sep. 5, 2003

Transformation of nonnormally distributed measures of condom use frequency can be problematic for researchers. Distributions are typically dichotomized. We used data collected from 483 university undergraduates in an anonymous, cross-sectional survey to illustrate the value of a screening analysis before dichotomization. Inconsistent condom users were compared to consistent users with respect to 14 measures. Subsequently, inconsistent users were compared to those who never used condoms with regard to the same 14 measures. Findings suggest that a screening analysis is a potentially important aspect of analyzing distributions that assess frequency of condom use.

KEY WORDS: Condom use; sexual behavior; sexually transmitted infections; dichotomy.

INTRODUCTION

Sexual transmission of the human immunodeficiency virus (HIV) poses a substantial threat to the health and well-being of people worldwide. Although condom use provides substantial protection against HIV, persons at risk of transmission or acquisition often report that condoms are used infrequently or not at all (Cates 2001; Steiner *et al.*, 1999). Indeed, a substantial level of effort has been dedicated to identifying correlates (in cross-sectional studies) or predictors (in prospective studies) of condom use among members of various populations. A recent review article on this topic delineated findings from 121 empirical studies (Sheeran *et al.*, 1999). Unfortunately, the many of these studies may have treated the dependent measure (i.e., condom use frequency) inappropriately.

Studies assessing condom use as a dependent variable can be plagued by multiple design issues and statistical problems that lead to Type 1 and Type 2 research errors (Crosby 1998; Crosby *et al.*, 2002). One very important, and commonly neglected, issue in these studies is the transformation of nonnormally distributed data. Assessments of how frequently people use condoms classically display a "U-shaped" (or nearly U-shaped) distribution. These distributions contain a preponderance of consistent users (i.e., 100% use) and "never" users (i.e., 0% use), with the remainder of the values falling between 1% use and 99% use (inconsistent users).

Because these distributions clearly violate assumptions necessary for parametric analyses, researchers have typically created one of two dichotomies: "consistent versus inconsistent" or "any use versus never." The difference between these two dichotomies is determined by whether inconsistent users are grouped with the consistent users or with "never" users. However, little is known about inconsistent users (Peterman *et al.*, 2000; Pinkerton and

[1]Rollins School of Public Health, Department of Behavioral Sciences and Health Education, Atlanta, Georgia.
[2]Emory Center for AIDS Research, Atlanta, Georgia.
[3]Rural Center for AIDS/STD Prevention, Bloomington, Indiana.
[4]The Kinsey Institute for Research in Sex, Gender, and Reproduction, Bloomington, Indiana.
[5]Department of Applied Health Science at Indiana University, Bloomington, Indiana.
[6]Gender Studies at Indiana University, Bloomington, Indiana.
[7]Correspondence should be directed 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, Georgia (e-mail: rcrosby@sph.emory.edu).

1090-7165/04/0300-0099/0 © 2004 Plenum Publishing Corporation

Abramson, 1996) and a clear rationale for grouping them with either consistent users or "never" users has not been provided in the literature.

Creating a dichotomy can be quite useful because this allows researchers to transform nonnormally distributed variables and to apply widely known and well-understood analytic methods such as contingency table analysis and logistic regression. However, selecting a cut-point for a dichotomy in the absence of empirical guidance is clearly problematic. Indeed, in cases where inconsistent users are statistically distinct from consistent and "never" users, creating a dichotomy is not justifiable; instead, a trichotomy or logarithmic transformation may be necessary.

A potentially useful solution to the question of how to handle inconsistent users when forming a dichotomy is to perform a screening analysis designed to establish an empirical basis for the grouping decision. A screening analysis simply informs analytic decision making when researchers are faced with critical decisions such as the transformation of outcome variables. Applied to the dependent variable of condom use, this analysis provides an empirical basis for subsequent decisions related to transforming a classic U-shaped distribution. Such an analysis may be a critically important step before analyzing condom use distributions that comprise dependent variables. Accordingly, this article demonstrates the utility of using a screening analysis to empirically guide the transformation of nonnormally distributed measures of condom use frequency. We chose to demonstrate this method using data from persons 18–24 years of age.

METHODS

Suggested Screening Analysis

A screening analysis can be based on bivariate associations. Thus, p values are the only test statistic required. The selected correlates or predictors (in this case we have arbitrarily selected 14 measures to serve as an example) are each subjected to two tests. The first test compares inconsistent users (those using condoms, but not on a consistent basis) to consistent users (those using condoms for every act of penile–vaginal sex). The second test compares inconsistent users to "never" users (those who did not use condoms for any act of penile–vaginal sex). These tests provide an empirical basis for subsequent decisions about the analytic treatment of the distribution representing frequency of condom use. Because these tests are not part of the statistical analyses that will ultimately be applied to the distribution, the use of multiple, bivariate, tests does not inflate the risk of Type 1 error.

Study Sample

From September 2001 through April 2002 research assistants enrolled 483 university undergraduates in an anonymous, cross-sectional survey. For analyses, we selected only those undergraduates who self-identified as heterosexual, were 18–24 years of age, were never married, and reported having penile–vaginal sex in the last 3 months ($n = 330$). The Institutional Review Board at Indiana University approved the study protocol.

Measures

Outcome Measure

Students provided information regarding frequency of penile–vaginal intercourse in the last 3 months. Frequency of condom use for the same time period was assessed. For students reporting at least one episode of penile–vaginal sex ($n = 347$), the latter measure was divided by the former to create a proportional measure of condom use.

Correlates

Because this analysis was designed only to exemplify the potential value of a screening analysis, our selection of correlates was atheoretical. In addition to race, gender, and age, we assessed several correlates typically employed in studies of condom use (DiClemente and Crosby, 2003; Santelli *et al.*, 1999; Warner and Hatcher, 1999). Specifically, we expected that greater levels of motivation, confidence and instruction in condom use, perceived effectiveness of condoms, number of sex partners, and nonuse of hormonal contraception (Warner and Hatcher, 1999) would be associated with more frequent condom use. In addition, based on previous studies (Foxman *et al.*, 1998, 2001) we anticipated that engaging in oral or anal sex could represent expanded sexual repertoires and therefore might be associated with less condom use.

Data Analysis

Associations between continuous correlates and the outcome measures were assessed by *t* tests. Associations between dichotomous correlates and the outcomes were assessed by chi-square tests. Significance was defined by alpha <.05. These bivariate tests allowed us to compare significant and nonsignificant associations for each correlate in the context of two different comparisons (i.e., inconsistent users vs. consistent users and inconsistent users vs. "never" users).

RESULTS

Average age was 19.4 years (standard deviation = 1.4). Eighty-eight percent self-identified as White, 6.1% self-identified as Black, and remaining students self-identified as other minorities. Fifty-six percent were female. Frequency of condom use was characterized by the classic U-shaped distribution (Fig. 1). Indeed, the distribution appears to have a U-shape without curves on the bottom corners. The distribution can therefore be viewed as consisting of "thirds." About one-third of the students (36.7%) never used condoms; one-third (33.0%) used condoms, but not on a consistent basis; and about one-third (30.3%) reported using condoms for every act of penile–vaginal intercourse in the last 3 months. In Fig. 1 the superimposed normal curve illustrates that the obtained scores are about as far from normal as is mathematically possible.

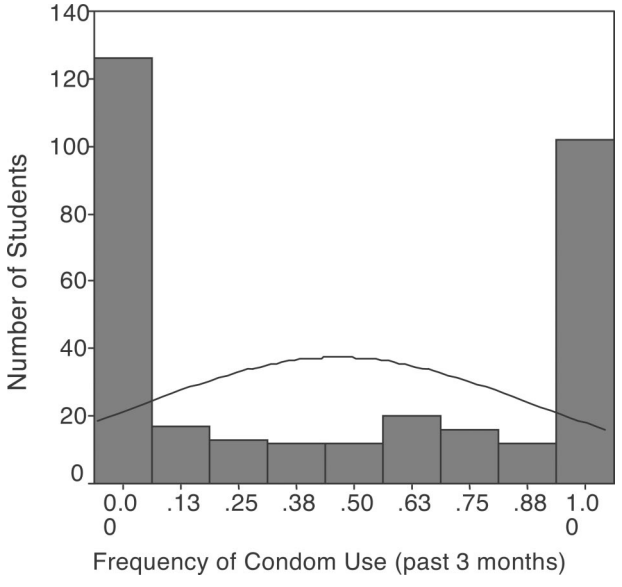

Fig. 1. Distribution of condom use frequency.

Table I. Comparison of Bivariate Significance for Correlates of Inconsistent Condom Use Versus Consistent Condom Use (Outcome A, $n = 209$) and Inconsistent Condom Use Versus Never Using Condoms (Outcome B, $n = 230$)[a]

	Outcome A: Inconsistent[b] vs. consistent[c]	Outcome B: Inconsistent[a] vs. never[d]
Continuous correlates	*p*	*p*
Age	.91	.43
Motivation to use condoms	.0001	.0001
Partner(s) motivation to use condoms	.0001	.0001
Confident about using condoms correctly	.19	**.009**
Believe condoms prevent STD/HIV	.99	**.03**
Believe condoms prevent pregnancy	.24	.14
Number of sex partners (last 3 months)	.36	.57
Frequency of penile–vaginal sex (last 3 months)	**.003**	.12
Dichotomous correlates		
Race	.12	.26
Gender	.87	.88
Self-reported oral sex (last 3 months)	.03	.002
Self-reported anal sex (last 3 months)	.24	.85
Ever received instruction about correct condom use	.22	**.046**
Relied on hormonal contraception (last 3 months)	**.003**	.24

[a] Bold signifies discrepant findings between the two outcome measures.
[b] $n = 109$.
[c] $n = 100$.
[d] $n = 121$.

Table I displays the *p* values representing the bivariate analyses that made up the screening analysis. In the table, comparisons between inconsistent users and consistent users is shown as Outcome A. Comparisons between inconsistent users and "never" users are shown as Outcome B. Of note, although these data are provided only as a means of demonstrating the potential value of a screening analysis, all significant associations shown in the table were in the anticipated direction.

Lack of Association

Students' age, beliefs about condom effectiveness for preventing pregnancy, number of sex

partners, race (dichotomized as minority vs. nonminority), gender, and self-reported anal sex were not significantly associated with either of the two outcomes.

Differential Association

Inconsistent users differed from consistent users with respect to frequency of sex and reliance on hormonal contraception. Conversely, inconsistent users differed from "never" users regarding confidence about using condoms correctly, beliefs about condom effectiveness against sexually transmitted diseases and HIV, and whether students had received condom use instruction.

Significant Associations for Both Comparisons

Inconsistent users differed from consistent users *and* from "never" users with respect to three correlates. Less motivation, less perceived partner motivation, and engaging in oral sex were associated with greater odds of inconsistent compared to consistent use *and* with greater odds of "never" use compared to inconsistent use.

DISCUSSION

The findings from this exercise suggest that the rationale for grouping inconsistent users with either consistent or "never" users should be a function of a screening analysis for the selected correlates. In the absence of such an analysis, arbitrarily grouping inconsistent users with either consistent or "never" users may lead to Type 1 and Type 2 errors. Instead, researchers should consider performing a screening analysis (as demonstrated here) to empirically justify their dichotomy. Findings from this procedure guide the grouping decision and set the stage for bivariate analyses designed to identify candidate predictor variables for a regression model. On a cautionary note, researchers should avoid using the screening test as way of pretesting predictor variables (i.e., looking for the optimal strength of association between any given predictor and the outcome variable) used in bivariate or multivariate hypothesis testing.

When *lack of association* (described in Results) is found, deciding how to group inconsistent users should be made on the basis of the research question or another established criterion. However, if *differential association* is found, the grouping decision becomes empirically driven. Thus, when inconsistent users differ from consistent users, but not from "never" users, the dichotomy should be inconsistent and "never" users compared to consistent users. Alternatively, when inconsistent users differ from "never" users, but not from consistent users, the dichotomy should be inconsistent and consistent users compared to "never" users.

Finally, when *significant associations for both comparisons* are found, a dichotomy may not be justified. In these scenarios, inconsistent users are quantitatively distinct from consistent users and from "never" users. Under these circumstances, researchers could create a trichotomy (analyzed using multinomial logistic regression) or perform logarithmic transformation of the original distribution (analyzed using linear regression). The exception to this would be cases where the researcher also tests observed bivariate associations for strength (a measure such as t for continuous predictor variables or a prevalence ratio for discrete predictor variables may be suitable). If one significant association is clearly stronger than the other, this information is certainly an important aspect in justifying a dichotomy.

Upon inspection of Table I, it becomes evident that selection of hypothesized predictors for a multivariate model should precede the decision to form a dichotomy. Furthermore, it becomes evident that creating a dichotomy in the absence of a screening analysis may indeed yield erroneous multivariate findings given that inconsistent users may be grouped inappropriately with consistent or with "never" users. For example, a researcher who hypothesized that frequency of sex and reliance on hormonal contraception would predict condom use is justified (based on the screening analysis) in creating a dichotomy that groups the inconsistent users with the "never" users (as the predictor variables do not distinguish between inconsistent users and "never" users). Using the same logic, the researcher would not be justified in creating a dichotomy that groups inconsistent users with consistent users (as the predictor variables discriminate between these two groups, which would subsequently become "one side" of the dichotomy). Alternatively, a researcher who hypothesized that participants' confidence in their ability to use condoms correctly and belief that condoms prevent STD/HIV would predict condom use is justified in grouping the inconsistent users with the consistent users rather than "never" users.

Although our demonstration of the suggested screening analysis was based on a relatively small convenience sample, the exercise supports an important methodological point. We showed that a simple screening analysis (based on bivariate associations) is a potentially important aspect of investigating research hypotheses that employ condom use frequency as a dependent variable. Of note, a screening analysis cannot entirely dictate analytic decision making. For example, the dichotomy justified by one predictor variable may not be justified (or even be counterindicated) by another predictor variable planned for the same multivariate model. Such discrepancies naturally necessitate cautious decision making before a transformation decision is reached. One objective of such deliberations is to avoid incorrectly grouping inconsistent users with one of the two remaining groups (i.e., consistent or "never" users). An incorrect decision would occur when a given predictor variable discriminates between two groups that were combined to form "one side" of a dichotomy (doing so would dilute the association and predispose the analysis to Type 2 error).

REFERENCES

Cates, W. (2001).The NIH condom report: The glass is 90% full. *Family Planning Perspectives*, *33,* 231–233.

Crosby, R. A. (1998). Condom use as a dependent variable: Measurement issues relevant to HIV prevention programs. *AIDS Education and Prevention*, *10,* 448–457.

Crosby, R. A., DiClemente, R. J., Holtgrave, D. R., and Wingood, G. M. (2002). Design, measurement, and analytic considerations for testing hypotheses relative to condom effectiveness against non-viral STIs. *Sexually Transmitted Infections*, *78,* 228–231.

DiClemente, R. J., and Crosby, R. A. (2003). Sexually transmitted diseases among adolescents: Risk ractors, antecedents, and prevention strategies. In G. R. Adams, and M. Berzonsky (Eds.), *Blackwell Handbook of adolescence* (pp. 573–605). Oxford: Blackwell.

Foxman, B., Aral, S. O., and Holmes, K. K. (1998). Heterosexual repertoire is associated with same-sex experience. *Sexually Transmitted Diseases*, *25,* 232–236.

Foxman, B., Aral, S. O., and Holmes, K. K. (2001). Interrelationships among douching practices, risky sexual practices, and history of self-reported sexually transmitted diseases in an urban population. *Sexually Transmitted Diseases*, *25,* 90–99.

Peterman, T. A., Lin, L. S., Newman, D. R., Kamb, M. L., Fishbein, M., Bolan, G., Zenilman, J., Douglas, J. M., Rogers, J., and Malotte, K. (2000). Does measured behavior reflect STD risk? An analysis of data from a randomized controlled behavioral intervention study. *Sexually Transmitted Diseases*, *27,* 446–451

Pinkerton, S. D., and Abramson, P. R. (1996). Occasional condom use and HIV risk reduction. *Journal of Acquired Immunodeficiency Syndromes*, *13,* 456–460.

Santelli, J. S., DiClemente, R. J., Miller, K. S., and Kirby, D. (1999). Sexually transmitted diseases, unintended pregnancy, and adolescent health promotion. *Adolescent Medicine State of the Art Reviews*, *10,* 87–108.

Sheeran, P., Abraham, C., and Orbell, S. (1999). Psychosocial correlates of heterosexual condom use: A meta-analysis. *Psychological Bulletin*, *125,* 90–132.

Steiner, M. J., Cates, W., and Warner, L. (1999). The real problem with male condoms is nonuse. *Sexually Transmitted Disease*, *26,* 459–462.

Warner, D. L., and Hatcher, R. A. (1999). *Male condoms*. In R. A. Hatcher, J. Trussell, F. Stewart, *et al.* (Eds.), *Contraceptive technology*, 17th ed. (pp. 325–352). New York: Irvington.