Sex Education Programs for Schools Still in Question

A Commentary on Meta-Analysis

Stan E. Weed, PhD

The meta-analysis reported in the three *Guide to Community Preventive Services* papers published in this issue of the *American Journal of Preventive Medicine*[1–3] targets two basic prevention strategies—comprehensive risk reduction (CRR) and abstinence education—which have the common purpose of lowering adolescent sexual risk behavior. Meta-analyses were conducted for each strategy on key behavioral and biological outcomes. This study did a commendable job of identifying the available studies of sexual risk-reduction interventions and evaluating them according to criteria for study quality. The Community Preventive Services Task Force (Task Force) also developed a rigorous category system to create appropriate dependent and independent variables for the analysis, and employed a sophisticated analytic strategy. They are to be commended for undertaking the study. Following are comments on the impact of heterogeneity on the study's conclusions, as well as comments on the authors' interpretation of study findings with regard to the public health impact of CRR and abstinence education.

Heterogeneity of the Comprehensive Risk Reduction Database and the Analytic Results

An important concern in meta-analytic studies is homogeneity: the extent to which there is adequate similarity among the interventions studied and consistency within the statistical results produced. Homogeneity is what allows the meta-analytic result to be interpretable and to inform the questions underpinning the study. Heterogeneity—the lack of homogeneity—is a common problem with meta-analysis and is measured by the I^2 statistic.[4–6] The CRR interventions represented in this study were very heterogeneous, comprising a wide variety of settings, populations, and pedagogies, with approximately 46% occurring within school-based programs and populations and 54% within community settings/populations—including programs in sexually transmitted diseases (STD) clinics, youth shelters, juvenile detention, or drug treatment facilities; community-based service-learning programs; multicomponent youth development programs; and programs for parents in low-income housing. By contrast, the majority (78%) of the abstinence education interventions occurred in school classroom settings with a general population of youth.

The high heterogeneity among comprehensive risk reduction *interventions* would be less problematic if the CRR *results* had been homogeneous. However, for seven of the ten CRR outcomes, the meta-analytic result produced an I^2 greater than 50%, indicating "notable heterogeneity," and the I^2 for five of these outcomes was above the recommended benchmark for "severe heterogeneity" of 56%[5]: sexual activity=67.67%, unprotected sexual activity=56.39%, use of protection=76.25%, condom use=78.26%, and use of both condoms and contraceptives=72.24% (Table 3[1]).

Scholars of meta-analysis have argued that such high heterogeneity in an estimate of overall effect can render the result meaningless at that level of aggregation.[4] The CRR result for sexual activity illustrates this problem. The I^2 of 68% indicates severe heterogeneity, and of the 35 programs represented, roughly one half (n=16, or 46%) had a neutral or negative effect (Figure 2[1]). Of these, 12 (or 34%) had negative effects. Although the authors conclude that CRR programs produced an across-the-board reduction in teen sexual activity of 12%,[1] the actual chances of finding a CRR program that reduces sexual activity appear to be about 50/50. The meta-analysis does not help end-users of sex education programs identify which half of the CRR programs are effective for this outcome, and even appears to mislead them by suggesting that all CRR programs will produce a reduction in sexual activity of at least 12% when, in fact, one third may actually increase sexual activity.

In cases of such high heterogeneity, scholars recommend that the causes of this inconsistency should be identified and reduced by examining subgroups with the "effect sizes . . . separated,"[4] and that conclusions should be drawn at that subgroup level.[4,6,7] The meta-analysis

From the Institute for Research and Evaluation, Salt Lake City, Utah

Address correspondence to: Stan E. Weed, PhD, Institute for Research and Evaluation, 6068 S. Jordan Canal Road, Salt Lake City UT 84129. E-mail: weedstan@aol.com.

0749-3797/$36.00

doi: 10.1016/j.amepre.2011.11.004

made an effort to do this by examining several demographic and program characteristics. Among these subgroups, evidence of differences was reported only for gender, and only on two outcomes, resulting in a conclusion in the Recommendations Statement that "these interventions may be somewhat more effective for boys than for girls."[3] However, study data produced by the Community Guide staff (data not shown) showed statistical evidence that CRR interventions are less effective in school settings than community settings.[8] These data show two outcomes (condom use and pregnancy) where the effect sizes for school programs were significantly worse than for community programs ($p<0.05$) and one where that difference approached significance (protection, $p<0.08$). (There were no outcomes where the effect sizes were significantly better in the school setting than the community setting.)

For all three of these outcomes, the effects for school-based programs had CIs that were not significant, whereas those for community-based effects were significant. This pattern was also true for the school-based sexually transmitted infection (STI) effect. It should be noted that the STI effect for school-based programs (OR=0.71, CI=0.24, 2.10) was the product of only two studies/programs with small sample sizes.[9,10] It should also be noted that the school effect for pregnancy was in the unfavorable direction, suggesting an increase in teen pregnancy (school-based OR=1.28, community-based OR=0.63, $p<0.04$).[8]

This evidence for school versus community differences appears to be more substantial than the evidence for gender differences in CRR effectiveness.[1] It appears to support a conclusion that CRR programs are less effective within school settings/populations than community settings/populations and constitutes a lack of evidence for the study's conclusion that CRR programs are effective within school settings/populations.[3]

Although the report of the meta-analysis acknowledges a heterogeneity problem for the CRR results, it does not appear to take this into account in the language of its conclusions and recommendations: "the evidence supports a conclusion that comprehensive risk reduction interventions are applicable across a range of populations and settings."[3] It is interesting that the study authors determined that the abstinence education results were "inconclusive because of inconsistencies in their effects on the outcomes examined"[1] when the CRR results show more evidence of inconsistent effects (seven outcomes with I^2 greater than 50%) than the abstinence education results (two outcomes with I^2 greater than 50%), albeit of another kind (Tables 3 and 6[1]). This high degree of heterogeneity in the CRR results also calls for inconclusiveness, not with regard to their validity but to their interpretation. Rather than drawing conclusions about the broad CRR category, it seems more appropriate to acknowledge an inability to draw general conclusions at this level of aggregation and to recommend further research to identify sources of heterogeneity and characteristics of effective programs.

The Question of "Dual" Comprehensive Risk Reduction Benefits

With regard to the public health impact of CRR versus abstinence education interventions, the study asserts that:

> Because CRR offers benefits both to adolescents who abstain from sex and to those who are sexually active, the overall public health impact for CRR is expected to be greater than that for abstinence education when the intervention effects on sexual activity are similar.[1]

This assumes that CRR programs are effective at promoting both abstinence and condom use. However, the benefits of any specific CRR program would be superior to the benefits of an effective abstinence education program only if that CRR program were affecting *both* behaviors. If it increased teen condom use but not abstinence, it would not offer a benefit that was superior to an effective abstinence program, because abstinence provides better protection for teens than condom use.

Thus, CRR should be viewed as a superior protective strategy only if it produces *both* outcomes *on the same population, within the same program*. Without the occurrence of this "dual effect" *within* CRR programs, decision makers (school administrators, youth advocates, parents) are left to choose between programs that do one thing or the other, that is, reduce sexual activity or increase condom use. Unfortunately, meta-analytic statistical methods are not suited to an empirical test of whether the CRR strategy has been effective at achieving both of these outcomes *within* CRR programs.

However, a study-by-study tally can be informative. Of the 83 CRR study arms, five provided evidence of significant dual effects. This represented three of the 62 CRR studies/programs. Although not all CRR studies measured effects on both outcomes (only 24 of the 62), we are still left with only three of the 62 CRR studies providing *evidence* of effectiveness at increasing both rates of teen abstinence and condom use. This is 5%, or about what would be expected by chance, and does not appear to constitute sufficient evidence for the assertion that the CRR strategy offers this dual effect and therefore provides greater public health benefits than the abstinence education strategy.

Summary

The researchers established a valid database founded on quality outcome studies and rigorous study methods. However, the high statistical heterogeneity of the comprehensive

risk reduction results raises concerns about the conclusion that the CRR strategy is uniformly effective, especially given the heterogeneity of the pool of interventions. The results were inconsistent across studies, indicating that some programs are working and some are not, but without identifying which of the many diverse types of CRR programs in the study are effective. Results from this meta-analysis made public elsewhere show a lack of evidence of effectiveness by school-based CRR at increasing rates of teen condom use and reducing pregnancy and STIs.[8] This is an important finding because schools are seen as a crucial conduit for sex education.

In addition, the lack of evidence for a dual CRR effect on abstinence and condom use within the same program does not support a conclusion that comprehensive risk reduction programs actually provide a public health benefit superior to effective abstinence education programs. Further research is needed to identify characteristics of effective CRR programs and to test for the existence of a dual CRR effect. The inconclusive abstinence education findings from the meta-analysis also underscore the need for more rigorous outcome studies of abstinence programs. It is my hope that the concerns articulated here will be addressed by future research so that group-based sex education can be implemented with greater success.

The author has consulted with Irene Ericksen of The Institute for Research and Evaluation in the preparation of this commentary.

No financial disclosures were reported by the author of this paper.

References

1. Chin HB, Sipe TA, Elder R, et al. The effectiveness of group-based comprehensive risk-reduction and abstinence education interventions to prevent or reduce the risk of adolescent pregnancy, human immunodeficiency virus, and sexually transmitted infections: two systematic reviews for the Guide to Community Preventive Services. Am J Prev Med 2012;42(3):272–294.
2. Sipe TA, Chin HB, Elder R, et al. Methods for conducting Community Guide systematic reviews of evidence on effectiveness and economic efficiency of group-based behavioral interventions to prevent adolescent pregnancy, human immunodeficiency virus, and other sexually transmitted infections: comprehensive risk reduction and abstinence education. Am J Prev Med 2012;42(3):295–303.
3. Community Preventive Services Task Force. Recommendations for group-based behavioral interventions to prevent adolescent pregnancy, human immunodeficiency virus, and other sexually transmitted infections: comprehensive risk reduction and abstinence education. Am J Prev Med 2012;42(3):304–307.
4. Sharpe D. Of apples and oranges, file drawers and garbage: why validity issues in meta-analysis will not go away. Clin Psychol Rev 1997;17(8):881–901.
5. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–58.
6. Moayyedi P. Meta-analysis: can we mix apples and oranges? Am J Gastroenterol 2004;99(12):2297–301.
7. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. Br Med J 1994;309(6965):1351–5.
8. Community Guide Staff. Univariate Analysis_CRR_031909. Guide to Community Preventive Services, CDC, 2009.
9. Smith MU, Dane FC, Archer ME, Devereaux RS, Katner HP. Students Together against Negative Decisions (STAND): evaluation of a school-based sexual risk reduction intervention in the rural south. AIDS Educ Prev 2000;12(1):49–70.
10. Kirby D, Korpi M, Adivi C, Weissman J. An impact evaluation of project SNAPP: an AIDS and pregnancy prevention middle school program. AIDS Educ Prev 1997;9(1S):44–61.