Distribution Agreement

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this dissertation. I retain all ownership rights to the copyright of the dissertation. I also retain the right to use in future works (such as articles or books) all or part of this dissertation.

Signature:

_____   _____
Wayne D. Johnson, Jr.   Date

HIV Prevention Research for Men Who Have Sex with Men:
Meta-Analysis, Intraclass Correlation, and Transformation
Between Count and Dichotomous Outcomes

By

Wayne D. Johnson, Jr.
Doctor of Philosophy
Epidemiology

_____
William M. McClellan, M.D., M.P.H.
Advisor

_____
W. Dana Flanders, M.D., D.Sc.
Committee Member

_____
Michael Goodman, M.D., M.P.H.
Committee Member

_____
Andrew N. Hill, Ph.D.
Committee Member

_____
David R. Holtgrave, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

_____
Date

HIV Prevention Research for Men Who Have Sex with Men:
Meta-Analysis, Intraclass Correlation, and Transformation
Between Count and Dichotomous Outcomes

By

Wayne D. Johnson, Jr.
M.S.P.H., University of South Carolina, 1987
B.A., University of South Carolina, 1980
B.S., University of South Carolina, 1980

Advisor: William M. McClellan, M.D., M.P.H.

An Abstract of
A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements of the degree of Doctor of Philosophy
in Epidemiology
2009

Abstract

HIV Prevention Research for Men Who Have Sex with Men: Meta-Analysis, Intraclass

Correlation, and Transformation Between Count and Dichotomous Outcomes

By Wayne D. Johnson, Jr.

This dissertation describes three methodological and content-based studies related
to meta-analysis of HIV prevention research for men who have sex with men (MSM).
First, we compared ANOVA models for individually- and group-randomized trials to
derive the factors necessary to account for intraclass correlation (ICC) in three classic
designs. For the simplest design, the factor is $(n-1)VIF / (n-VIF)$ where n is the number
of participants per condition, and VIF is the variance inflation factor $1+(m-1)ICC$.
Simulations confirmed our correction factors for both additive and multiplicative models.

Second, we used regression and the method of moments to identify candidate
formulas for transformation between proportions (summarizing dichotomous data), and
means and variances (summarizing count data). Best empirical results were obtained
from regression models predicting the proportions as a function of the mean and variance
(and vice versa), or by the method of moments assuming a negative binomial distribution.

Third, we applied these results in a meta-analysis of behavioral HIV prevention
for MSM. We found 54 interventions with 16,224 participants. The 38 interventions that
were compared to minimal or no HIV prevention controls reduced unprotected sex by
27% (95% confidence interval [CI] = 15% to 37%). The other 16 interventions reduced
unprotected sex by 17% compared to standard or other HIV prevention interventions (CI
= 5% to 27%). Our methodological work permitted a robust conclusion that behavioral
prevention reduces self-reported unprotected sex among MSM.

HIV Prevention Research for Men Who Have Sex with Men:
Meta-Analysis, Intraclass Correlation, and Transformation
Between Count and Dichotomous Outcomes

By

Wayne D. Johnson, Jr.
M.S.P.H., University of South Carolina, 1987
B.A., University of South Carolina, 1980
B.S., University of South Carolina, 1980

Advisor: William M. McClellan, M.D., M.P.H.

A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements of the degree of Doctor of Philosophy
in Epidemiology
2009

Acknowledgments

I thank my committee for their encouragement, advice, and insights. I am particularly grateful to my advisor Bill McClellan for his patience and guidance through the dissertation process, and to Dana Flanders and Andrew Hill for many hours of direction with statistical issues.

Table of Contents

List of Tables

List of Figures

1. Overview

The condition that came to be known as Acquired Immune Deficiency Syndrome (AIDS) was initially recognized in 1981 when clusters of *Pneumocystis* pneumonia (CDC 1981a), Kaposi's sarcoma (CDC 1981b), and other opportunistic infections (Gottlieb 1981) were identified among previously healthy gay men in Los Angeles and New York City. Laboratory tests revealed a common mechanism of profound cellular immune dysfunction (Masur 1981).

Before a cause was identified, the epidemiological similarities to hepatitis B led researchers and gay men to suspect a viral agent transmissible by blood and semen (Francis 1983). Within 2 years of recognition of the condition, CDC published recommendations to reduce anal sex without condoms and number of sex partners (CDC 1983). Ethnographic network mapping revealed that some 40 of the first 100 known cases were included in a network of sexual contacts among men whose residences ranged from New York to California to Texas and Florida (Auerbach 1984). Identification of the virus that came to be known as Human Immunodeficiency Virus (HIV) was announced almost simultaneously by laboratories in Paris (Barré-Sinoussi 1983), Bethesda (Gallo 1983), and San Francisco (Levy 1984).

Now approaching the fourth decade of the epidemic, a vaccine remains elusive, treatments are expensive, have debilitating side effects and do not cure, and men who have sex with men (MSM) continue to account for the largest proportion of new infections in the United States and much of the industrialized world. Behavioral prevention remains critical to the effort to minimize transmission.

Effects of behavioral interventions to reduce risk of sexual transmission of HIV

among MSM have been evaluated in several randomized trials and strong quasi-experimental studies. There is a need to optimize the usefulness and interpretability of these results through quantitative synthesis. Substantive questions to be addressed relate to the overall effectiveness of interventions, and how effects differ according to characteristics of interventions, populations, and research methods.

Several methodological challenges must also be addressed in order to combine and compare data across a broad range of study designs, interventions, and populations. Critical among these are accounting for the unit of assignment to treatment status in community-level studies and transformation between count and dichotomous outcomes.

This dissertation addresses methodological and content issues regarding summary, stratified, and regression analyses of effects of HIV prevention interventions for MSM. The Background chapter examines the state of the epidemic, previous reviews, and methodological concerns. The Objectives chapter identifies three issues of content and methodology that have been raised and not yet answered by the current literature. The Methods chapter describes epidemiological, statistical, and behavioral science-based procedures for addressing the three issues identified in the Objectives chapter. The Results chapter gives a brief overview of the findings of three studies. Manuscripts reporting the studies themselves are provided as appendices at the end of the dissertation as described below.

After the Results section, the Discussion draws conclusions from the three studies taken together, discusses strengths and limitations that may not be addressed in the individual studies, and describes an agenda for further research. A chapter of References is then presented which includes the citations from all previous chapters as well as the

two methodological manuscripts and the published journal article.

Finally the research studies themselves are presented as appendices. Each study contains detailed methods, results, and conclusions. The two methodological studies are presented in a form compatible with submission to a scientific journal. Those two studies are titled "Correcting the Variance of the Intervention Effect in Group-Randomized Trials When Only the Variance Appropriate to an Individually-Randomized Trial Is Available" (included as Appendix A) and "Transformation Between Count and Dichotomous Outcomes for Meta-Analysis: Estimating Risk Ratios from Means and Rate Ratios from Proportions" (included as Appendix B).

The meta-analysis itself has already been published and is titled "HIV Intervention Research for Men Who Have Sex with Men: a 7-Year Update," which is included as Appendix C. This work has also subsequently been updated and re-formatted as a Cochrane Review, which is included as Appendix D and titled "Behavioral Interventions to Reduce Risk for Sexual Transmission of HIV Among Men Who Have Sex with Men." Both the journal article and the Cochrane Review are included with permission as appendices at the end of this dissertation.

2. Background

This chapter includes first a review of literature about the epidemiology and prevention of HIV/AIDS, followed by reviews of the literature motivating each of the three papers that constitute the results section of this dissertation.

2.1 Epidemiology and Prevention of HIV/AIDS

Three laboratories almost simultaneously reported identification of the virus that came to be known as HIV (Barré-Sinoussi 1983, Gallo 1983, Levy 1984). The agent was a retrovirus, encoded as RNA and capable of reverse coding itself as DNA into the host genome. Retroviruses in humans were rare, although two that caused leukemia had been identified in the Caribbean and southern Japan and had similar transmission patterns. The etiologic agent was thus recognized as either a truly new pathogen or one that was emerging from a previously isolated environmental or geographic base. Retroviruses have coevolved with several species including primates, horses, sheep, and cats. Molecular epidemiology now suggests that HIV entered human populations from other primates on several different occasions, sometimes from slaughter and consumption of primates as food. Retroviruses from chimpanzees apparently evolved within humans into the various clades of HIV-1 (Gao 1999), and others from sooty mangabees to become HIV-2 (Clavel 1986), a similar but somewhat less aggressive virus that is so far still generally limited to West Africa.

A blood test for antibodies to HIV became available in 1985 (Sarngadharan 1984) introducing new opportunities for screening donated blood, public health surveillance, more sensitive outcomes in epidemiological studies (HIV infection rather than AIDS), and voluntary testing and counseling of people who considered themselves at risk. With

this opportunity came bad news: Infection rates were vastly greater than had been imagined, with perhaps 100 times as many infected as symptomatic. Even worse news followed: Progression to disease appeared to approach 100% among those infected (Lui 1988). The time from infection to symptoms of AIDS was estimated at 7.8 years, 95% CI 4.2 to 15 years. Hopes for quick development of a vaccine evaporated when genetic analysis revealed that HIV mutations within a single infected person rivaled the annual global diversity of influenza A (Korber 2001).

2.1.1 Biology of HIV infection

In initial infection, Langerhans or dendritic cells perform their normal role in the immune response by carrying HIV to T cells that carry the CD4 molecule, a gateway for HIV's entry to the cell (Dalgleish 1984, Klatzmann 1984). The T cells then carry the virus to lymphoid tissue, again the normal function where antigens are processed for mounting an immune response. But immune activation induces massive HIV replication mediated by cytokines and aberrant cell signaling due to an interaction between the viral envelope and cellular receptors (Kinter 2000). The immune system is caught in a cycle of partial viral control and accelerated viral replication, leading to depletion of the CD4+ T cell population and eventual destruction of the immune system (Fauci 1996, Fauci 2003).

Along with the CD4 receptor, the CCR5 chemokine coreceptor is also critical to initial uptake of the virus; people who are homozygous for a deletion in the gene that produces this molecule (e.g., about 1% of white populations) are extremely resistant to HIV infection, while heterozygotes show some resistance and delayed disease progression (Paxton 1998, O'Brien 2000). Initial infection apparently requires CCR5 to transfer a non-syncytium inducing ("R5") form of the virus (D'Souza 1996). Within the

infected person, a new subset of syncytium-inducing ("X4") virus then evolves that relies

on the CXCR4 chemokine for transfer between cells (Scarlatti 1997).

Examination of highly exposed but persistently seronegative (HEPS) female sex

workers in Kenya (Rowland-Jones 1998) and northern Thailand (Sriwanantha 2001)

demonstrated other mechanisms of resistance to infection that were not related to CCR5.

In northern Thailand, the infectibility of CD4+ cells, the suppressive capacity of CD8+

cells, and production of β-chemokines did not differ from those of non-HIV-exposed

individuals. Instead, the distinctive property of blood samples from HEPS individuals

was production of a soluble activity that suppressed post-integrated HIV-1 replication

(Butera 2001). This activity was produced only when monocytes and CD4+ T cells were

cultured together, suggesting that the protection these women experience is due to

restriction of HIV transfer from infected macrophages to CD4+ cells. In Kenya, CD8+ T-

cell responses were observed but no serum antibodies. Some sex workers became

susceptible after leaving the trade, suggesting that continuous exposure is necessary for

immune protection (Kaul 2001).

2.1.2 Medical treatment and prevention efforts

A new era of treatment for AIDS and HIV disease was introduced with the

availability of highly active antiretroviral therapies (ART or HAART) (Fauci 2003).

These approaches attack various targets in the viral replication cycle, from reverse

transcriptase inhibitors (e.g., AZT) to protease inhibitors (Flexner 1998) to the most

recent class of virus-cell fusion inhibitors (Burton 2003). Combinations of these therapies

lead to remarkable clinical improvement and reductions of plasma viral loads to

undetectable levels. But even after 3 years of such treatment, viral levels rebound from

reservoirs in certain components of the immune system if therapy is not maintained (Chun 1999).

Effective vaccines remain elusive (McMichael 2003). Even with full-blown chronic infection, the humoral and cell-mediated immune responses are only partly successful in controlling viral replication, so a natural model for true immunity is lacking. Superinfection of HIV-infected but otherwise healthy people apparently does occur, although it is not clear how often (Atfeld 2002, Goulder 2002), suggesting that a single vaccination followed by sterilizing immunity (as with smallpox or measles) may be difficult to achieve.

2.1.3 Risk of transmission through sexual contact

Early in the epidemic, receptive anal intercourse without a condom was identified as the riskiest sexual activity for MSM (Darrow 1987). Two-thirds (240) of 492 MSM who had been enrolled in 1978-80 in studies of hepatitis B in San Francisco had seroconverted for HIV by 1985. The strongest risk factors for infection were receptive anal intercourse with ejaculation by nonsteady partners, many sexual partners per month, and other indicators of high levels of sexual activity. The vast majority (95%) reported engaging in receptive anal intercourse since joining the cohort; 69% of these men subsequently became infected. In contrast, of the 18 men who had not had receptive anal intercourse, only 4 (22%) seroconverted (odds ratio = 7.9, 95% CI = 2.5-24.6). A clear dose-response was also evident: compared to a 29% seroconversion rate among those who reported no anal exposure to ejaculate from nonsteady partners, infections occurred in 53% of those who reported exposures to a few such partners (<1 per month) (OR = 2.8), 78% of those reporting 1-2 such partners per month (OR = 8.5), and 85% of those

reporting 3 or more such partners per month (OR = 13.6). The overall number of male

partners and events of bleeding during or immediately after intercourse were also

associated with seroconversion, but other sexual activities were not strongly predictive

after accounting for receptive anal intercourse.

In the Multicenter AIDS Cohort Study, 95 (3.8%) of 2507 initially HIV-negative

MSM seroconverted during six months follow-up (Kingsley 1987). The only significant

risk factor for new HIV infection was receptive anal intercourse, with a risk ratio of 3 for

one partner and 18 for five or more partners.

More recently the risk of HIV transmission per episode of unprotected receptive

anal intercourse (URA) has been estimated as 0.82% (95% CI, 0.24% to 2.76%) when the

partner was known to be HIV+ and 0.27% (95% CI, 0.06% to 0.49%) when partners of

positive or unknown status were included (Vittinghoff 1999). But this risk was strongly

heterogeneous: Nine men became infected after only one or two episodes of URA with

HIV+ or unknown serostatus partners. This prospective cohort study of 2,189 high-risk

homosexual and bisexual men was conducted in San Francisco, Denver, and Chicago in

1992-1994. The risk of infection through other sexual activities with HIV+ or unknown

serostatus was estimated to be much lower: 0.06% (95% CI, 0.02% to 0.19%) per contact

for unprotected insertive anal intercourse and 0.04% (95% CI, 0.01% to 0.17%) per

contact for unprotected receptive oral sex to ejaculation.

One factor implicated in heterogeneity of risk is variation of infectiousness of

infected partners. Mathematical models of the early epidemic in San Francisco suggested

that the risk of infection from a single episode of unprotected receptive anal intercourse

with a recently infected partner might be as high as 50%, while the risk with a partner in

the long asymptomatic phase was on the order of 1 in 1000 to 1 in 10,000 (Jacquez 1994, Koopman 1997). Such figures might suggest a prevention model based on reducing the rate of new partner acquisition. If upon infection, whether the event is recognized or not, a participant has sex with no other partners until the highly infectious phase is past (e.g., six weeks), then the risk of his subsequently transmitting HIV to a new partner is greatly diminished. At the community level, the reproductive rate (the R_0) for new infections would fall below one, and the epidemic would end. However this may not be the typical pattern of human sexual activity. It could also be considered a drawback of such an approach that the benefit of reduced risk accrues not directly to those who comply but to their partners. While the risk of transmission through unprotected anal intercourse from an infected insertive partner to an uninfected receptive partner is less than 1%, the risk to the receptive partner is extremely heterogeneous, with almost a quarter of infections occurring with only a single exposure (Vittinghoff 1999).

A meta-analysis revealed that risk behavior typically decreases substantially after receipt of a positive HIV test; however risk behavior *increases* slightly after a negative result (Weinhart 1999). Although risk behavior decreases after receipt of a positive HIV test, it does not cease. In a US cohort of 66 MSM, 39% reported insertive UAI with an HIV-negative or unknown serostatus partner in the 6-month study period during which their seroconversion occurred. This proportion dropped to 2% during the first month after their first HIV-positive test, but rose again to 13% by nine months (Colfax 2002).

Mathematical modeling of partnership formation among young MSM in Amsterdam suggests that a surprisingly high proportion - 86% - of HIV transmission occurs within steady rather than casual partnerships (Xiridou 2003). Assuming that

highly active retroviral therapy (HAART) will dramatically reduce transmission by 75-99%, the authors recommend promotion of HIV testing (increasing coverage from 42% to 80%) and HAART administration (increasing from 70% to 85%) as prevention strategies.

Research suggests that disclosure is the norm, but 16% of HIV-seropositive gay or bisexual men reported unprotected sex without disclosure, and 3.2% reported insertive sex to ejaculation without disclosure in the 6 months prior to interview (Ciccarone 2003). In many circumstances, HIV-infected men may be assuming that their partners are also already infected, and this assumption may often be incorrect.

2.1.4 MSM still at high risk for HIV and STD transmission

Men who have sex with men (MSM) still constitute the largest group of new AIDS cases each year in much of the developed world. In the United States for example, among those for whom a risk category was identified, MSM (including MSM who inject drugs) accounted for 50% of people newly diagnosed with AIDS in 2001, and 64% of new diagnoses among men (CDC 2001a). This high representation of MSM was evident among all racial and ethnic groups, ranging from 51% of new diagnoses among African American men to 81% among white men. This high proportion of cases is borne by a group believed to constitute only 2% to 10% of the adult male population (Kinsey 1948, Binson 1995).

In the states with confidential HIV reporting, MSM accounted for 52% of all persons and 71% of men with newly reported HIV infections for whom an exposure category was reported in 2001 (CDC 2001b). These percentages are essentially the same as the cumulative proportions for the entire epidemic (53% and 71% respectively). Back-

calculation methods indicate not only that the majority (53%) of all new HIV infections in the United States in 2006 were among MSM, but also that the rate of new infections is increasing among MSM while decreasing among injection drug users and high risk heterosexuals (Hall 2008).

Increases in rates of unprotected anal intercourse (UAI) among gay men have been reported in large urban centers. In a longitudinal study, the proportion of gay men in San Francisco reporting UAI nearly doubled (from 20% to 39%), with the most dramatic increases among men 26-29 years old (Ekstrand 1999). The self-reported increases in UAI have been validated by observed increases in the rates of rectal gonorrhea in this population (CDC 1999). The San Francisco Department of Health has also reported increases in the estimated incidence of HIV infection among MSM between 1997 and 1999 (McFarland 1999).

By the mid-1990s, unprotected anal intercourse with nonprimary partners was uncommon among MSM in West Hollywood, California (Crepaz 2000). But the prevalence of UAI with primary partners was greater than with nonprimary partners, and was most prevalent among men younger than 25 years of age. Thus young MSM may be at risk for HIV through their sexual risk behavior with primary partners.

But the turn of the century has seen substantial increases in STDs including syphilis, gonorrhea, and herpes simplex virus type 2 among MSM, suggesting that new increases in HIV incidence may be forthcoming (Wolitski 2001). Challenges in addressing this next wave of infections include the tedium of consistent condom use, new cultural and technological milieus through which HIV may spread but for which behavioral interventions have not been developed or tested, lethargy in biomedical

prevention while treatment is emphasized, and cultural failure to address stigmatized topics including morality, mortality, drug use, and sex, particularly between men (Gross 2003b).

2.1.5 HIV prevention needs among MSM

Behavioral strategies that MSM have used to reduce risk in the past may need reinforcement or refinement, particularly among groups where HIV prevalence and risk behavior remain high, such as young MSM and MSM of color (Valleroy 2000, Jones 2008). Significant challenges in vaccine development add to the urgency of identifying and promoting effective behavioral intervention strategies (Stott 1999, Lancet Editorial 2000). Recent increases in sexual risk behavior and HIV transmission demand a closer look at the available data regarding effective behavioral interventions for MSM.

MSM are confronted with the AIDS epidemic as one of numerous health and social issues including violence, depression, suicide, drug and alcohol addiction, and social stigma (Paul 2002, Gross 2003a). Interactions among these problems may have a syndemic effect, each exacerbating the others (Stall 2003).

In the United States, black MSM have been disproportionately affected by HIV/AIDS (CDC 2001c). The higher infection rates are not explained by higher reported rates of unprotected sex, but may be due to underreporting of risk behavior, increased prevalence of HIV among sexual contacts, increased infectiousness of sexual partners due to coinfection with other STI, or increased susceptibility (Malebranche 2003).

Baseline data from the EXPLORE study, an intervention project currently underway, indicate that among 4295 HIV-negative MSM in 6 US cities, 48% reported unprotected receptive anal sex and 55% reported unprotected insertive anal sex in the

previous 6 months (Koblin 2003). Almost 13% reported having had a specific STD in the 6 months before they enrolled in the study. Men with one primary partner and men with multiple partners were slightly more likely to report unprotected sex than men with one nonprimary partner. Unprotected sex was more common with drug and alcohol use.

While MSM are at high risk and account for a large proportion of new HIV infections each year, prevention funding for program and research efforts directed toward this group have been less than proportional. The Centers for Disease Control and Prevention observe that only 15% of federal AIDS prevention resources explicitly target this population (Valdiserri 2002).

2.1.6 Related risk factors

In Calgary, Alberta, MSM who contracted syphilis were more likely to be older, to be coinfected with HIV, and to report heavy alcohol use (versus injection drug use) than heterosexuals with syphilis (Jayaraman 2003). These MSM most often used the Internet and bars or bathhouses to meet sexual partners.

Transgendered people with or without sexual reassignment may be at elevated risk if their sexual partners are at risk. This group faces many additional challenges and should not be neglected in HIV prevention efforts (Fee 2003).

Women who have sex with women may also be at risk, largely due to elevated risk among men whom they also have sex with, or among people they share drugs with (Friedman 2003). They may be more likely to have sex with or share needles with MSM or to have been institutionalized or homeless.

2.1.7 Condoms for HIV prevention

In 2000, the National Institute of Allergy and Infectious Diseases (NIAID)

convened a workshop to answer the question: "What is the scientific evidence on the effectiveness of latex male condom use to prevent STD transmission during vaginal intercourse?" (NIAID 2001). STDs considered included HIV infection, gonorrhea and chlamydia (including pelvic inflammatory disease), syphilis, chancroid, trichomoniasis, genital herpes, and genital human papillomavirus (HPV) infection and associated diseases (cervical dysplasia, cervical cancer, and genital warts). The panel found the published literature to be inadequate to definitively answer the broad question regarding STDs in general, mainly because few studies employed the optimal prospective design to assess the effectiveness of condoms in preventing infection.

However the panel did conclude, based on a meta-analysis (Davis and Weller 1999) of published studies that the strongest published data documenting effectiveness of the male condom were for prevention of HIV transmission. In that meta-analysis, 12 cohort samples of serodiscordant heterosexual couples yielded a consistent HIV incidence of 0.9 per 100 person-years (95% CI, 0.4 to 1.8) among those who always used condoms. This value was compared to 6.8 per 100 person-years (95% CI, 4.4 to 10.1) for male-to-female transmission and 5.9 per 100 person-years for female-to-male transmission in 11 cohort samples of those who never used condoms. (No estimate was available concerning male-to-male transmission.) The overall protection afforded by condoms was estimated as 87%, with a range from 60% to 96% depending on the incidence among condom non-users.

Regardless of sexual orientation, 18% (95% CI, 16% to 19%) of sexually active adults in the United States report using a condom at last intercourse with partners in an ongoing sexual relationship, and 43% (95% CI, 36% to 49%) with other sexual partners

(Anderson 2003). Outside of ongoing relationships, condom use at last intercourse was more common among those who were not married (56%), those with 2 or more partners in the past year (61%), and those reporting sex with a stranger (65%). No increase across time was observed from 1996 to 2000.

In a study in Alabama, 1996-1999, sexually active African American adolescent females were recruited from medical clinics and high schools to examine the effect of condoms on preventing infection with chlamydia, gonorrhea, or trichomonas (Crosby 2003). Among those who reported using condoms every time they had sex, 17.8% tested positive for one of these infections at 6-month follow-up, compared to 30.0% among those who reported less than 100% condom use, a 40.7% reduction. The difference did not appear to be attributable to number of partners or frequency of intercourse.

Several follow-up and case-control studies among female sex workers (FSW) in Southeast Asia indicate high efficacy of consistent condom use in reducing risk of HIV transmission. In a follow-up study in Thailand, HIV incidence was reduced by 92% among FSW who reported using condoms every time they had sex with clients (1.8% per year) *vs.* others (23.8% per year) (Kilmarx 1998). Another follow-up study of STD incidence examined pharyngeal gonorrhea among FSW in Singapore (Wong 1999). Those who reported inconsistent condom use for oral sex were 17 times more likely than consistent condom users to develop pharyngeal gonorrhea during 6 months follow-up. Only 2.6% of FSW who always used condoms for oral sex contracted pharyngeal gonorrhea during the study period compared to 11.7% of others, a 78% reduction in risk.

Finally in Southeast Asia, policy interventions have also shown favorable effects on HIV prevention. In Thailand, the 100% Condom Program was an administrative

requirement beginning in about 1990 that condoms be used for all sex in commercial sex environments, plus a mass advertising campaign. According to national surveillance, STD cases (which were already decreasing) among men at government clinics decreased by 86% from 1989 to 1994 (Rojanapithayakorn 1996). HIV incidence decreased from 4% to 2.7% among new army conscripts. HIV prevalence came to a plateau of about 30% among "direct" FSW (those whose primary work activity is to sell sex) and 10% among "indirect" (those who sell sex as a secondary activity in conjunction with waitressing, bartending, etc), and condom use for commercial sex increased from less than 20% to about 95%. During the same years, the number of registered FSW decreased by 23%, and the proportion of FSW who worked directly decreased from 57% to 33% (Hanenberg 1998). Men began to patronize sexual services less frequently, prices rose, and a higher proportion of FSW came from outside Thailand.

Less information is available concerning the effectiveness of condoms for preventing HIV transmission in sex between men. In a mathematical model, condom use for receptive anal sex reduced HIV transmission by only one-third, from a risk of 0.27% to 0.18% per contact with HIV+ or unknown serostatus partners (Vittinghoff 1999). The authors note that reported condom failure explained much of the remaining risk, that the receptive partner may sometimes not be aware of condom failure, and that respondents may sometimes report unprotected episodes as protected due to stigmatization. Other explanations for these infections may include mistaken perceptions that partners were HIV-negative (episodes with partners perceived to be negative were excluded from the model) or unusually long time from infection to seroconversion (with higher risk behavior occurring before the 6-month recall period). Three of the 52 men who

seroconverted during the study reported sex only with men believed to be HIV-negative, so some misclassification appears likely, whether of behaviors, partners' serostatus, or timing of infection.

Breakage and slippage have been identified as reasons for condom failure. A randomized controlled trial revealed no significant difference in failure rates for thicker condoms (2.3% breakage or slippage) versus standard (2.5%) when used by male couples (Golombok 2001).

2.1.8 Alternatives to use of male condoms

The effectiveness of alternatives to use of male condoms also appears to have been more thoroughly studied in high-risk heterosexual populations than among MSM. Female condoms were found to be a highly effective alternative to male condoms in an RCT among FSW in Thailand (Fontanet 1998). While condom use was very high in both groups, the proportion of acts for which condoms were not used (2.1%) among FSW assigned to use female condoms when their clients refused male condoms was 22% lower than in the group assigned to use only male condoms (2.7%). Parallel to the difference in behavior, STD incidence was 24% lower in the group who had the option of using female condoms than in those restricted to male condoms.

A case-control study showed that exposures due to condom breakage or slippage were less common when clients of FSW wore two or more condoms, one inside another (Rugpao 1997). Condoms broke in 1.8% of sex acts where only a single condom was used, compared to 0.2% of acts with 2 condoms and none with 3 or more condoms. The practice appeared to be quite common, being reported for just over 50% of sex acts, and was credited for a decrease in breakage reports from 6% in 1992 to 1% in 1995.

Use of microbicidal gels containing nonoxynol-9 was found to be ineffective in reducing HIV-1 infection among FSW in Africa and Southeast Asia, where 59 (16%) of 376 women randomly assigned to use nonoxynol-9 became infected, compared to 45 (12%) of 389 placebo users (Van Damme 2002). In a case-control study (Fihn 1996), urinary tract infections were 3 times more common among women whose partners used spermicide-coated condoms than among women whose partners did not use condoms. Authors of both studies speculated that the spermicide may irritate exposed vaginal or cervical mucosa, thus increasing risk of HIV or STD transmission.

Mixed results have been obtained in trials of enhanced treatment for STDs as an approach for HIV prevention in developing countries (Wasserheit 1992). In Mwanza, Tanzania, improved syndromic treatment of STDs plus a behavioral intervention targeting condom use and prompt treatment of STD symptoms led to 58% fewer HIV infections in six intervention communities than in six matched comparison communities (Grosskurth 1995). However a substantial proportion of the difference in incidence appears to be attributable to baseline differences in prevalence among the 12 communities. The impact of behavior change in response to the behavioral aspect of the intervention was also not addressed.

In Rakai, Uganda, a second major trial took STD treatment a step further: Treatments for several STDs (gonorrhea, chlamydia, syphilis, trichomoniasis, bacterial vaginosis) were administered to all consenting adults in treatment communities (Wawer 1999). Treatments were directly observed and were delivered door to door. Presumptive mass treatment was considered justified because of the general high prevalence of STDs. Participants in control communities received vitamins and anti-helminthics. This study

found no difference in HIV incidence between STD intervention and control communities. Subsequent mathematical modeling suggests that sexual risk reduction associated with restored civil stability may have been the critical factor in reducing Uganda's HIV epidemic, which had once been among the most severe in Africa (Korenromp 2002).

2.1.9 Risk factors for progression to HIV disease and AIDS

Numerous behavioral and biologic factors have been noted as potential cofactors of progression from asymptomatic HIV infection to more active HIV disease including AIDS or to death. In a San Francisco cohort of 370 MSM with well-characterized seroconversion dates, decreased survival time was most strongly associated with weekly hallucinogen use (relative hazard (RH) = 2.6, 95% CI=1.6 to 4.3) and receptive anal intercourse with ejaculation (RH = 1.45, 95% CI = 1.02 to 2.04) (Vittinghoff 2001). Progression to AIDS diagnosis was associated with the same variables (RH for weekly hallucinogen use = 2.59, 95% CI=1.56 to 4.28) (RH for receptive anal intercourse with ejaculation = 1.4, 95% CI = 1.0 to 2.0) as well as with weekly cocaine use (RH = 1.5, 95% CI = 1.1 to 2.1). No association was found with age at seroconversion; use of other drugs, alcohol, or tobacco; post-seroconversion STDs including gonorrhea, syphilis, or HBV infection; or other sexual practices.

Some genetic characteristics appear to be associated with improved survival time. HLA-B (human lymphocyte antigen type B) alleles can be divided into two mutually exclusive groups based on the expression of the molecular HLA-Bw4 and HLA-Bw6 epitopes. Homozygosity for HLA-Bw4 is associated with profound suppression of HIV-1 viremia, maintenance of normal CD4 T-cell count, and delayed time to AIDS symptoms

(Flores-Villanueva 2001). A deletion allele of the CCR5 (or CKR5) structural gene also results in reduced host infectibility (Smith 1997, Dean 1996).

2.1.10 Correlates of sexual risk

At the Pittsburgh site of the Multicenter AIDS Cohort Study (MACS) in the late 1980s, HIV positive MSM were somewhat *less* likely to have engaged in high-risk sexual activity, particularly unprotected insertive anal sex, within the past 6 months than HIV-negative MSM (Robins 1997). Regardless of serostatus, MSM reporting unprotected insertive anal sex were younger, less educated, had less psychological distress and greater feelings of mastery, employed fewer behavioral coping strategies, drank more alcohol and used more amyl nitrate (poppers).

Gay-identified men in Chicago were more likely than bisexual men to engage in receptive sex, including unprotected anal sex (Stokes 1997). Gay men were more likely to have had a steady male partner or lover, but total number of partners did not differ significantly between the two groups. Bisexual men were more self-homophobic and saw other people as less accepting of same-sex activity. Thus different intervention strategies may be necessary for non-gay identified MSM.

In six US cities, 1999-2001, unprotected anal sex (both receptive and insertive) with HIV-positive and status-unknown partners correlated with enjoyment of unprotected receptive anal sex and with low scores for self-efficacy for safer sex, communication skills, and social norms (Chesney 2003). These psychosocial correlates of risk may be useful for intervention development.

Condoms must be available if they are to be used. In Massachusetts, the Board of Education recommended in 1991 that school districts consider providing condoms in

secondary schools; 10% of school districts with high schools decided to provide condoms (Blake 2003). Instruction in prevention of HIV, STD, and pregnancy was generally enhanced along with provision of condoms. In 1995, the proportion of students who had ever had sexual intercourse was reduced by 14% in districts that provided condoms compared to other districts (42% vs. 49%); among those who were sexually active, the proportion who did not use condoms during their most recent sexual intercourse was reduced by 36% (28% vs. 44%). Thus provision of condoms and risk reduction programs promoted condom use among high school students without increasing sexual activity.

2.2 Previous reviews of HIV prevention for MSM

Given the urgency of HIV prevention for MSM, rigorous evaluations of the effects of interventions for MSM were at first slow to accumulate. Until the late 1980s, the most pertinent information concerning risk reduction was obtained from longitudinal surveys (Stall 1988) rather than from controlled trials of interventions. Those studies demonstrated that, despite profound behavioral risk reduction resulting from formal and informal intervention efforts, many MSM continued to be at substantial risk for HIV transmission. However none of these interventions were being tested against comparison conditions. (Some researchers distinguish between two terms to refer to the non-experimental groups: "control" in studies with random assignment, and "comparison" in studies with non-random assignment. For simplicity we will henceforth refer to "control or comparison" groups as "comparison" groups.) Since then, several reviews have identified studies of HIV prevention interventions and suggested hypotheses to explain variations in effectiveness.

In 1991, Peterson, Ostrow, and McKirnan (Peterson 1991) identified five studies

of HIV prevention for MSM; two of these evaluated experimental interventions against comparison conditions. Kelly (1989) found less unprotected sex and more condom use for anal sex after a 12-session small group intervention offering AIDS information, skills training, and motivation. After a skills and information intervention (compared to information only), Valdiserri and colleagues (1989) found an increase in number of partners with whom condoms were used during insertive (but not receptive) anal sex.

In 1992, Fisher and Fisher (Fisher 1992) identified 48 studies of interventions involving psychological and/or educational elements designed to modify an outcome relevant to AIDS-risk reduction and subjected to formal statistical evaluation. Thirteen (27%) of these focused on MSM, three of which evaluated interventions against comparison conditions. In addition to the Kelly and Valdiserri studies cited above, Coates (1989) found a reduction in number of sex partners among HIV-positive men after a stress reduction program. The reviewers concluded that interventions were more likely to reduce risk behavior if they provided not only information about AIDS transmission, but also the motivation to change risky behavior and particularly the skills to apply risk-reduction strategies.

Hays & Peterson (1994) identified eight studies evaluating certain aspects of HIV prevention for gay and bisexual men in metropolitan cities. Only two of these measured changes in HIV risk behavior in experimental intervention conditions vs. comparison, control, or other intervention conditions. In addition to the Valdiserri study noted above, this review described the Mpowerment study, outcomes of which were later to be published by Kegeles, Hays, and Coates, 1996. After conducting a peer-led community-level program including outreach, small groups, and a publicity campaign, this study

found decreases in the proportion of young men reporting unprotected anal intercourse both with boyfriends and with nonprimary partners.

In 1994, Choi and Coates (Choi 1994) undertook a comprehensive critical review in which they identified 77 scientific reports on HIV prevention studies. Only 10 (13%) of these, six of which included comparison groups, focused on MSM. In addition to the three studies published in 1989, and preliminary reports concerning the Mpowerment study as described above, Choi and Coates identified Kelly 1992, and a preliminary report concerning Kelly 1997, in both of which gay men in communities receiving bar-based opinion-leader interventions reported reductions in unprotected anal sex contrasted against gay men in communities not receiving the intervention. The reviewers concluded that small groups and community interventions can produce at least short-term change and that brief *skills training* can maintain change. The authors pointed out that no studies had yet presented outcome data specific to MSM who are young, who are not white, or who do not identify themselves as gay.

In 1995, Holtgrave *et al.* (Holtgrave 1995) observed that HIV prevention programs can reduce risk behavior if they have "sufficient resources, intensity, and cultural competency and are based on behavioral and social science theory and past research." There were too few intervention studies available at the time to perform a quantitative review, but their identification of characteristics of effective interventions led to several testable hypotheses regarding intervention components, target populations, and research design that could be useful in subsequent meta-analyses.

Oakley *et al.* (Oakley 1995) restricted their scope to studies with comparison groups, pre- and post-intervention data, and information on all targeted outcomes. They

found 18 studies that met these criteria, of which only 4 (22%) concerned MSM; two of these (Kelly 1989 and Kelly 1992 had been identified by previous reviews. Of the two studies newly identified, both included multiple intervention arms and a comparison arm. Rosser found no statistically significant differences in effects on safer sex among four diverse intervention conditions and a comparison arm. Tudiver found a greater decrease in unsafe anal sex after a single-session group intervention led by a peer volunteer than after a four-session group led by paid counselors, or among control participants. The reviewers recommended that HIV intervention research should apply stronger evaluation designs.

Kalichman, Carey, and Johnson reported a meta-analysis of cognitive and behavioral outcomes of HIV risk-reduction interventions across all populations in 1996 (Kalichman 1996). Only 2 (17%) of the 12 eligible studies were focused on MSM. A weighted least-squares regression suggested that intervention effects diminished as the time from intervention to follow-up progressed from 1 to 6 months.

We found four more reviews or commentaries from 1996 on the status of HIV prevention research for MSM. In a resource guide for use by HIV prevention Community Planning Groups (Middlestadt 1996), only 5 (15%) of a list of 33 articles focused on MSM. Oakley, Oliver, Peersman and Mauthner (Centre for the Evaluation of Health Promotion and Social Interventions) subsequently cast a more finely meshed methodological net in an update of their previous summary. Coates *et al.* (Coates 1996) noted that although substantial risk reduction has already taken place, continued and evolving behavioral prevention efforts are essential; efforts should be targeted to those at greatest risk for infection and transmission, including but not limited to MSM and people

who already have HIV; and the science of HIV prevention must be enhanced, including improved surveillance methods, vaccine development, and linkage from science to practice. Finally Graham Hart (Hart 1996) underscored the importance of quantifying the effectiveness of interventions on changing behavior in real world settings. Yet no new rigorous evaluations of interventions for MSM were available for inclusion in any of these reviews.

In a review of recent HIV prevention interventions for gay men, Kegeles and Hart (1998) identified three additional intervention studies with outcome data: (1) Gold and Rosenthal (1995), in which gay men who were assigned to keep a diary of their sexual activity and provide self-justifications for episodes of anal sex without a condom were less likely to report more than one unprotected episode than men who were assigned only to keep a diary; (2) Peterson *et al.* (1996), where African-American MSM randomly assigned to a triple-session intervention were less likely to report unprotected anal intercourse (UAI) than their counterparts assigned to a wait-list control; and (3) Choi *et al.* (1996), who found that Asian or Pacific Islander MSM who were assigned to a single-session, 3-hour intervention group reported fewer sex partners than API men who were assigned to a wait list control condition. This review also provided preliminary reports concerning several evaluation studies which had not yet presented outcome data. The reviewers noted that community-level interventions may be advantageous in that they can reach people who have not already decided that HIV prevention is important or that they are at risk. By contrast, the strength of individual and small-group approaches is that they can deliver a larger number of hours of intervention to each participant than can be delivered, on average, to each individual within a community setting. These reviewers

also noted a methodological concern that because many of these populations have received the attention of numerous prevention efforts in recent years, it may now be difficult to determine the true effect of newly introducing any given intervention.

A quantitative analysis of 9 intervention studies was presented in 2002 (Johnson 2002). Of approximately 100 studies which met inclusion criteria concerning research design and outcome measures, only 10 focused on MSM, 9 of which reported behavioral outcomes. A summary of effects of those interventions revealed a consistent significant reduction in unprotected sex when contrasted against comparison conditions.

Thus previous reviews of HIV prevention point to the need for a quantitative summary and analysis of prevention studies focusing on MSM. Past reviews have identified effective interventions and have suggested several hypotheses regarding intervention content (information, motivation, skills; resources, intensity, and cultural competency), delivery (community-level, small groups, hours of intervention), populations (by age, race/ethnicity, gay or non-gay identification), and measurement (evaluation designs, time from intervention to follow-up) to explain differences in effectiveness. However, a need remains for an up-to-date quantitative summary and analysis to integrate the lessons learned and to evaluate which preventive strategies prove most effective under what circumstances.

2.3 Accounting for the unit of assignment in group-randomized trials

Many health promotion interventions are designed for delivery to social groups that already exist before the intervention is implemented. Murray (1998) describes examples such as school-based smoking prevention programs, mass media campaigns to reduce delay time in seeking medical care during a heart attack, selection of heart-healthy

foods from restaurant menus, and training of bar staff to reduce service to patrons who appear intoxicated.

In such cases, the health or behavioral outcome of interest occurs at the level of individuals (smoking, seeking medical care, selecting items from menu, drinking or serving drinks), but the interventions can only or best be delivered to larger groups (schools or classrooms, media markets, restaurants or communities, bars). Group-randomized trials are often the best method for evaluating effects of such interventions.

In a group-randomized trial, the unit of assignment to treatment condition is the group, but behavioral or health outcome data is typically gathered at the individual level. Thus g schools may be randomized to receive an experimental program to reduce smoking and g schools to a control condition (thus $c = 2$ conditions), but m students (members) within each school may be surveyed concerning behavior (or their cotinine levels may be examined etc.)

Several researchers and methodologists have emphasized that because the unit of assignment in such studies is the group rather than the individual, the error variance against which the intervention effect should be compared is determined primarily by variance among these groups rather than among individuals. An analysis that fails to take this into account and instead treats the data as if each person had been individually randomized to treatment or control conditions will generally overstate the precision of the results (and thus understate the variance), often badly. A classic description of the problem was:

Randomization by cluster accompanied by an analysis appropriate to

randomization by individual is an exercise in self-deception, however, and should

be discouraged.

Cornfield (1978), pp 101-102.

Unfortunately, research teams may not have the means or may not be aware of the

need to conduct an analysis that accounts for groups as units of assignment. Funding for

health promotion research may be so scarce that studies must be limited to only a few

clusters, in which case researchers may resort to individual-level analyses to "enhance"

power. It is not uncommon for a group-randomized trial to include only a single cluster

per treatment condition, in which case cluster is entirely confounded with treatment

condition, and no variance remains to be allocated to variation among clusters by the

usual methods.

Compatible procedures for variance estimation are necessary if group-randomized

studies are to be combined and compared in meta-analyses with studies where individuals

are the unit of assignment. Better estimates of variance may be valuable to program

planners and policy makers who seek to make the best use of limited budgets for health

promotion efforts. For such purposes, a formula is needed to estimate the variance that

would have been obtained from an analysis appropriate to randomization by cluster given

only the results of an analysis appropriate to randomization by individual.

2.4 Transformation between different types of data for meta-analysis

A method to transform results between dichotomous and count measures is

needed to permit comprehensive analysis of randomized controlled trials that use the two

types of outcomes. A method has previously been described for transformation between

means and dichotomies assuming an underlying normal (or technically logistic)

distribution (Hasselblad and Hedges, 1995; Chinn, 2000; Johnson, Semaan, Hedges *et al*.,

2002). This method assumes that the preferred outcome is the odds ratio or the

standardized mean difference (which we define in the manuscript in Appendix A).

Although valid odds ratios and standardized mean differences can be estimated, risk

ratios or rate ratios may be preferred for randomized trials because these metrics are

directly interpretable as the ratio of either the proportion reporting any occasions or the

mean number of occasions between the two treatment conditions, and risk odds ratios can

be misleading when events are common (Deeks, Higgins, Altman *et al*., 2008; Deeks,

2002; Altman and Deeks, 1998; Katz, 2006; Deeks, 1998; Bracken and Sinclair, 1998).

Thus new methods are needed to transform directly between risk ratios and rate ratios.

3. Objectives

In the studies included as appendices to this dissertation, we first present two papers addressing specific statistical and methodological challenges that we encountered as we undertook a meta-analysis of HIV prevention research for MSM. In Study I (included as Appendix A), we derive and test correction factors to account for the unit of assignment in group-randomized trials when the only available data are collapsed across groups. In Study II (Appendix B), we derive and test equations to permit transformation between outcome measures when only count outcomes or dichotomous outcomes are available.

We then present a meta-analysis to identify, describe and summarize rigorous HIV prevention studies for MSM (Study III, Appendix C). We also include a Cochrane Review (Appendix D, not numbered as a study) in which the meta-analysis is further updated and re-formatted to meet criteria of the Cochrane Collaboration.

The studies are titled: 1) Correcting the Variance of the Intervention Effect in Group-Randomized Trials When Only the Variance Appropriate to an Individually-Randomized Trial Is Available, 2) Transformation Between Count and Dichotomous Outcomes for Meta-Analysis: Estimating Risk Ratios from Means and Rate Ratios from Proportions, and 3) HIV Intervention Research for Men Who Have Sex with Men: a 7-Year Update. The Cochrane Review is titled "Behavioral Interventions to Reduce Risk for Sexual Transmission of HIV Among Men Who Have Sex with Men."

4. Methods

This section describes the statistical and procedural methods used in each of the three studies (clustered assignment, transformations, and meta-analysis). Each study also includes its own methods section. Study III was published first and uses some terminology (Prevalence Ratio as opposed to Risk Ratio) and methods (assumption of a constant dispersion factor) that differ from Study II.

4.1 Methods for determining the factor necessary to adjust for clustered assignment

We used algebra to derive and compare the formulas for components of the ANOVA table for the model appropriate for individual-randomized trials, for example for the post-test only design:

$$Y_{i:l} = \mu + C_l + \varepsilon_{i:l} \tag{1}$$

with the ANOVA table for the model appropriate for cluster-randomized trials, again for example for the post-test only design:

$$Y_{i:k:l} = \mu + C_l + G_{k:l} + \varepsilon_{i:k:l} \tag{2}$$

We applied this process to three designs: the post-test only design, the nested cross-sectional design, and the nested cohort design. The proposed correction formulas were derived as the ratio of the variance of the mean of the intervention effect under model 2 to the variance under model 1.

Correction factors obtained by this process apply explicitly to linear models where the outcome is usually continuous. However, the most commonly available data were dichotomous (any unprotected sex vs. none during the recall period), the analytical method most frequently applied was logistic regression, and the metric most commonly used for presenting these results was the odds ratio. The sums of squares framework we

used to develop the correction factors under the linear model does not apply to logistic models. Therefore it remained to be shown whether the same correction factors would apply to logistic models. We generated simulated data to test the performance of the correction factors. To examine the validity of each formula, we compared the adjusted variance resulting from each of the three formulas in simulations where the correct variance can be shown. These procedures are described in greater detail within the manuscript.

We then applied the resulting factor to several examples of group-randomized trials to reduce unprotected sex among men who have sex with men. We examined the effects of correcting for ICC on each separate study as well as on meta-analyses of all studies under several different sets of assumptions. Further details are provided within the manuscript.

4.2 Methods for determining formulas to transform between count and dichotomous outcomes

We applied regression and the method of moments to derive several candidate formulas for estimating proportions given only means and variances, and for estimating means and variances given only proportions.

Quantities resulting from those formulas were then substituted into the usual equations for the risk ratio (for dichotomies) and the variance of the log risk ratio, as well as the rate ratio (for counts) and the variance of the log rate ratio, to obtain five series of estimated intervention effects and variances under varying sets of assumptions. The meta-analytical weight could then be estimated in the usual way as the reciprocal of the variance. We also applied the delta method assuming either a Poisson, geometric, or

negative binomial distribution to obtain an additional series of methods for estimating the variance of the intervention effect.

Then we compared the performance of the resulting candidate formulas on empirical data sets from behavioral HIV prevention research for men who have sex with men. Formulas yielding minimal mean squared error for the log of the intervention effect and the log of the weight were deemed to provide optimal performance. Further details are provided in the manuscript. Again note that we used the term *risk ratio* in study III to represent the same metric that we previously referred to as the *prevalence ratio* in study I. Either term is correct.

4.3 Methods for meta-analysis of HIV prevention research for MSM

The following section describes the methods used within the 2005 publication, "HIV intervention research for men who have sex with men: a 7-year update," (*AIDS Education and Prevention* 2005;17:568–589). Some of these methods were further refined and enhanced during the development of the other two manuscripts included in this dissertation. In particular, we found that the regression methods described in the "Transformations" manuscript (in Appendix B) provided improved precision over the assumption of a constant dispersion parameter *d* for the negative binomial distribution (on page 39 and following in this section). We also use the term "prevalence ratio" in the 2005 publication, and within this section 4.3, to identify the same concept that is identified as the "risk ratio" in the Transformations paper.

4.3.1 Search strategies and eligibility criteria

We systematically reviewed the HIV prevention literature to find studies measuring the effects of behavioral interventions for MSM (Semaan *et al*., 2002).

Resources included online databases (e.g., Medline, PsychInfo, PubMed, AIDSLine, Web of Science), reviews and other studies in the HIV prevention literature, expert recommendation, hand searches of selected journals, and manuscripts and unpublished reports submitted by researchers.

Keywords for electronic searches varied according to database. As an example, a Medline search in August 2004 for (AIDS/prevention & control [pc] or HIV infections/pc or sexually transmitted diseases/pc) yielded 24,143 citations. A search for (homosexuality or bisexuality or gay.mp or bisexual.mp or men who have sex with men.mp or seropositivity/psychology) yielded 13,262 citations, and a search for (randomization or intervention studies or program evaluation or random.mp or randomize.mp or randomized.mp or randomly.mp) yielded 292,874 citations. Most quasi-experimental studies included the terms "intervention studies" or "program evaluation." Of the 77 citations included in all three searches, 49 were potentially eligible trials or reviews of HIV prevention interventions. Review of these 49 led to identification of 21 trials that were eligible by the criteria described below.

The potentially eligible HIV prevention studies from all sources were then evaluated by criteria of outcomes measured and study design. We included only studies that measured intervention effects on behaviors understood to affect risk of HIV transmission (e.g., unprotected sex, condom use, number of partners) and biologic outcomes including incidence of infection by HIV or other STD. We defined unprotected sex as anal intercourse without a condom. Data concerning other sexual and drug use behaviors were not frequently available. Only three eligible studies reported biologic outcomes. Because unprotected anal sex is the most epidemiologically pertinent behavior

for MSM (O'Leary, DiClemente, & Aral 1997) and was available for all studies, we conducted analyses only for this outcome.

We excluded interventions that focused not on sexual transmission but on cognitive or affective outcomes such as distress associated with HIV testing, or health and coping for seropositive men (Perry, Fishman, Jacobsberg, Young, & Frances 1991; Chesney, Chambers, Taylor, Johnson, & Folkman 2003). We included only studies in which MSM constituted all or a substantial proportion of the study sample (e.g., HIV-seropositives) or were specifically targeted by the intervention. When other populations were included, we obtained outcome data for the MSM subset or reduced the study weight to reflect only the proportion who were MSM.

Acceptable study designs were randomized controlled trials and certain quasi-experimental designs. Quasi-experimental studies were required to include independent comparison groups assigned without bias, that is, without regard to volition, self-selection, need, or other baseline characteristics, and to include separate baseline data for the intervention and comparison groups. We requested supplemental information from authors when intervention effects, or the data necessary to calculate them, e.g., separate results by study arm, were not published.

4.3.2 Calculation of effect sizes

Because eligible studies used randomized or quasi-experimental designs, we chose rate ratios (RR) to estimate intervention effects for count measures and prevalence ratios (PR) for dichotomous measures (Greenland 1998, Deeks 1999). Note that we used the term *prevalence ratio* in study I to represent the same metric that we later refer to as the *risk ratio* in study III. Either term is correct.

For each study that reported count measures (number of episodes of or partners for unprotected sex), the rate ratio at follow-up was the ratio of the mean in the intervention group to the mean number in the comparison group. The natural logarithm of the rate ratio (LnRR) was then an estimate of the intervention effect. The reciprocal of the variance of LnRR (see Appendix) served as a measure of the weight of information provided by the study.

Similarly the prevalence ratio at follow-up was the ratio of the proportion of respondents reporting unprotected sex in the intervention group to the proportion in the comparison group. The natural logarithm of the prevalence ratio (LnPR) was an estimate of the intervention effect, and the reciprocal of the variance of LnPR (see Appendix) estimated the weight of the study. Rate ratios and prevalence ratios less than one represented a difference favoring the experimental intervention group.

When individual-level data were available, we used SAS Proc Genmod to estimate intervention effects adjusted for the baseline value of covariates such as the outcome variable, age, race/ethnicity, and serostatus. For count outcomes, we used the negative binomial distribution and the log link function and adjusted the scale for Pearson's chi-square divided by its degrees of freedom to estimate the rate ratio. For dichotomous outcomes, we used the binomial distribution and the log link function to estimate the prevalence ratio. Scale adjustment does not apply to dichotomous outcomes.

When individual-level data or adjusted statistics were not available, we adjusted for the baseline distribution of the outcome variable by subtracting the baseline LnRR or LnPR from the follow-up LnRR or LnPR. We used the lesser of the baseline and follow-

up weights for such studies. These decisions concerning baseline adjustment and weights were based on the available empirical examples with raw data.

In studies where communities are the unit of assignment to treatment, the variance of the intervention effect will be underestimated if intraclass correlation (ICC) is not accounted for (Murray 1998). We applied the adjustment factor developed in study I to reduce study weights where necessary to account for ICC. For small values of ICC, the adjustment factor is approximately equal to Donner's variance inflation factor (VIF): $1 + ICC \times (m - 1)$ where m is the number of subjects in each unit of assignment. We assumed an ICC of .005, the value observed in the one study for which ICC was published (Kelly, *et al*., 1997).

For studies that measured results at multiple follow-up times, we used data representing cumulative effects closest to 12 months after the intervention. We used outcome variables that did not distinguish between insertive and receptive sex, main and nonmain partners, or partners perceived to be seroconcordant *vs.* serodiscordant when such data were available. For studies from which the only available results were separated by insertive *vs.* receptive sex, or main *vs.* nonmain partners, we used the average point estimate and the average weight of the two measures to estimate the underlying combined effect (Johnson, Semaan, Hedges, Ramirez, Mullen, & Sogolow 2002b). When results were not available concerning main or seroconcordant partners, we accepted results concerning only nonmain, serodiscordant, or unknown serostatus partners. For studies that compared two or more experimental interventions against a single control group, we divided the control group into equal parts for comparison to each of the interventions.

4.3.3 Statistical analysis

We applied the standard procedures for meta-analysis to conduct summary, stratified, and regression analyses (Cooper & Hedges 1994). We conducted separate analyses using rate ratios and prevalence ratios. In order to include all eligible studies in each analysis, we substituted prevalence ratios for rate ratios in studies that measured only dichotomous outcomes, and vice versa in studies that measured only count outcomes. If the prevalence ratio is constant across cutpoints then the rate ratio equals the prevalence ratio. This assumption of constancy across cutpoints is analogous to the assumption used to justify transformation between log odds ratios and standardized mean differences in other meta-analyses (Hasselblad & Hedges 1995, Chinn 2000, Johnson *et al*., 2002b). This assumption appeared plausible based on the studies for which effects at multiple cutpoints were available. Variances (and therefore weights) however differ substantially between count and dichotomous outcomes. We used the following formulas to estimate the variance of LnRR when only dichotomous data were available, and the variance of LnPR when only count data were available. Specifically, when only dichotomous outcome data (e.g., sample sizes *m* and *n* and sample prevalences $\hat{p}$ and $\hat{q}$) were available we estimated the variance of LnRR as:

$$Vâr\left(Ln\hat{R}R\right) = \frac{\hat{d}\left(1-\hat{p}\right)^{\hat{d}}}{m\left[1-\left(1-\hat{p}\right)^{\hat{d}}\right]} + \frac{\hat{d}\left(1-\hat{q}\right)^{\hat{d}}}{n\left[1-\left(1-\hat{q}\right)^{\hat{d}}\right]}$$

For the pooled dispersion parameter $\hat{d}$ for the intervention and comparison groups we used the value 6.5, the weighted mean value of the dispersion parameters from 19 studies from which it could be directly estimated. The dispersion parameter is estimated as:

$$\hat{d} = \frac{V\hat{a}r[Z] - \hat{E}[Z]}{\hat{E}^2[Z]}$$

where $\hat{E}[Z]$ and $V\hat{a}r[Z]$ are the pooled sample mean and the pooled sample variance for all intervention and comparison groups in a study. This formula applies when the underlying count data are overdispersed relative to the Poisson distribution, that is, the variances are greater than the means.

When only count data (e.g., sample sizes m and n, sample means $\hat{E}[Y]$ and $\hat{E}[X]$, and sample variances $V\hat{a}r[Y]$ and $V\hat{a}r[X]$) were available we estimated the variance of LnPR as:

$$V\hat{a}r\left(Ln\hat{P}R\right) = \frac{\hat{f}^{\hat{j}}}{m\left(1 - \hat{f}^{\hat{j}}\right)} + \frac{\hat{g}^{\hat{k}}}{n\left(1 - \hat{g}^{\hat{k}}\right)}$$

where

$$\hat{f} = \frac{\hat{E}(Y)}{V\hat{a}r(Y)}$$

$$\hat{j} = \frac{\hat{E}^2(Y)}{V\hat{a}r(Y) - \hat{E}(Y)}$$

$$\hat{g} = \frac{\hat{E}(X)}{V\hat{a}r(X)}$$

$$\hat{k} = \frac{\hat{E}^2(X)}{V\hat{a}r(X) - \hat{E}(X)}$$

$\hat{E}(Y)$ = intervention sample mean

$V\hat{a}r(Y)$ = intervention sample variance

$\hat{E}(X)$= comparison sample mean

$V\hat{a}r(X)$ = comparison sample variance

This formula applies when the data are overdispersed, that is, the sample variances are greater than the sample means. Note that we later developed additional formulas for the Transformations paper (Study III); the methods we used in Study I, which was published in 2005, are not necessarily the same methods we want to recommend after Study III.

We present results separately for interventions contrasted against minimal to no HIV prevention control conditions and those contrasted against standard or other HIV prevention conditions. We defined minimal to no HIV comparison conditions as including no treatment, wait lists, lagged designs, counseling for emergencies only, passive display of materials in community settings, and several treatments not addressing sexual behavior (diet and exercise training, substance abuse treatment, health support groups, and medication adherence consultation). Standard or other comparison conditions included HIV prevention seminars, individual HIV prevention counseling and testing, HIV prevention videos, and keeping a diary of sexual activity in the context of HIV prevention.

We conducted both random and fixed-effects meta-analyses (Hedges & Vevea 1998). Because intervention effects were generally homogeneous, results of the two types of models were usually identical. When results did differ, point estimates from the fixed effects models were slightly (about 1%) more conservative. Therefore we present only results of fixed effects models. We used stratified analyses to examine subgroup effects according to intervention format (small group, individual, or community-level), and to summarize interventions for HIV-positive MSM. We applied the standard principles of

weighted meta-regressions (Cooper & Hedges 1994) to account for multiple study characteristics and to examine differences in effects according to exposure rates in community-level interventions.

We examined the potential effect of outlier studies by excluding each intervention effect one at a time and recalculating the summary effect. To investigate the possibility of publication bias, we examined a linear regression through the funnel plot of treatment effect on sample size (Macaskill, Walter, & Irwig 2001). To be concise, we present regressions, sensitivity to outliers, and analysis of publication bias only for rate ratios and not prevalence ratios.

5. Correcting the Variance of the Intervention Effect in Group-Randomized Trials when Only the Variance Appropriate to an Individually-Randomized Trial Is Available

See manuscript attached as Appendix A.

6. Transformations Between Count And Dichotomous Outcome Measures For Meta-Analysis: Estimating Risk Ratios from Means and Rate Ratios from Proportions

See manuscript attached as Appendix B.

7. Meta-Analysis of HIV Prevention Research for Men Who Have Sex with Men

See published article attached (with permission) as Appendix C.

Also see published Cochrane Review attached (with permission) as Appendix D.

8. Discussion

8.1 General conclusions

These three studies contribute to both the substantive and methodological understanding of meta-analysis and of HIV prevention research for men who have sex with men. In Study I we developed correction factors for the variance of the intervention effect in studies where the unit of assignment to treatment condition is groups rather than members. The correction factors were derived for linear models, where the outcome may be either continuous or dichotomous, and the intervention effect is measured as the difference between conditions. We showed by simulations that the correction factors also perform well in logistic models where the intervention effect is measured as the odds ratio. When ICC is small, the correction factors are approximately equal to the VIF, except for the case of the completely naïve model for the cohort design, in which case the correction factor is approximately $V\hat{I}F \times \left(1 - \hat{r}_{yy(m)}\right)$. This simplified approach also performed well in simulations.

Study II showed that regression or the method of moments can be used to estimate risk ratios given only means and variances, or rate ratios given only proportions. Compared to simply interchanging the risk ratio with the rate ratio, regression and the method of moments assuming a negative binomial distribution provided substantial improvements in estimates of the risk ratio. However, the same methods yielded no improvements in precision of estimates of rate ratios over simple substitution of risk ratios. Estimates of weight were greatly improved for both directions. It would be most beneficial for researchers to measure and report dichotomous unprotected sex as well as number of episodes for and number of partners for unprotected sex as three separate

outcomes. When the wider range of preferred information is not provided, methods such as regression and the method of moments may be useful.

Study III showed that behavioral interventions for MSM reduced episodes of or partners for unprotected sex by 27% compared with minimal or no intervention, and reduced the proportion of men reporting any unprotected sex by 16%. Study III also found that count outcomes such as number of episodes of or partners for unprotected sex may be more sensitive than dichotomous outcomes, which do not recognize even a very large decrease in an individual's risk unless unprotected sex is altogether eliminated. A reduction in number of occasions of unprotected sex may have an important impact on HIV transmission, particularly if the number of partners for unprotected sex and the density of unprotected sexual networks also decrease.

8.2 Strengths

Study I showed that appropriate post hoc corrections can be made to meta-analytical weights for studies where existing groups are the unit of assignment. Our derivations were based on additive models, but simulations showed that the correction factors worked well for multiplicative models as well in the range of correlations typical of intervention research in this field. In our empirical examples, the correction factors were often as great as 6 or 7 even at a modest ICC of .005.

Outside of meta-analysis, this correction process is also useful for individual studies where only one cluster has been assigned to each treatment condition. Confidence intervals and statistical significance can be presented for a range of assumed values of ICC from similar studies, for example zero (no correction), .005, and .030. Even when two or three groups have been assigned to each condition so that ICC can be estimated,

the estimate has a wide confidence interval and it may be prudent to use these formulas to consider a range of possible values for ICC.

Although group-randomized trials often have little power to detect a statistically significant intervention effect when ICC is properly accounted for, meta-analysis of several such studies can provide stable and valid estimates. If an external estimate or likely range of ICC is available, correction factors for linear and logistic models are straightforward and easy to apply. The VIF (or $\hat{VIF} \times \left(1 - \hat{r}_{yy(m)}\right)$ in the case of the completely naïve model for the cohort design) performs well in the range of correlations we examined. This correction can also be used to estimate variance even in the case of only one or a small number of clusters per treatment condition, which is common in community-level intervention studies, and where estimates of ICC are unstable or cannot be estimated.

Study II examined a wide range of candidate estimates (regression and several statistical distributions) for estimating proportions given means and variances, and vice versa. Nineteen eligible studies were identified including 79 comparisons. Methods to substantially reduce mean squared error from the default procedures were identified for the log risk ratio and its variance as well as for the variance of the log rate ratio.

Study III was a comprehensive meta-analysis of 54 HIV prevention interventions for MSM, summarizing a literature that has grown rapidly in recent years but without standardization of outcomes, designs, or intervention content. As we approach the fourth decade of the AIDS epidemic, MSM are still the population at greatest risk for HIV in the developed world. Intervention effects were homogeneous, consistent with the hypothesis that a consistent effect was occurring across studies. And while by our estimates only five

of the studies independently attained statistically significant results, meta-analysis across all studies and several subsets of studies showed that HIV prevention interventions for this population do consistently yield reductions in self-reported unprotected sex.

8.3 Limitations

In Study I, our derivation does not address matched or stratified designs, multiple waves of followup data collection, or analyses of followup data with adjustment for baseline conditions as a covariate. This correction process adjusts only the variance and associated statistics such standard error and meta-analytical weight. It cannot address concerns regarding the validity of the point estimate when only one or a few clusters have been randomized to each treatment condition. Such point estimates are subject to bias because an insufficient number of units have been randomized to account for different trajectories that might have occurred even in the absence of the intervention. This bias can be reduced by combining multiple studies within meta-analysis.

In Study II, empirical data did not closely fit any of three distributions presented (Poisson, geometric, negative binomial), nor did they fit two other distributions that we considered but did not present here (zero-inflated Poisson or zero-inflated geometric). The closest fit appeared to be the zeta distribution, but parameter estimates for that distribution were lower than the minimum necessary to permit estimation of means and variances. The zero-inflated negative binomial distribution may have yielded a good statistical fit to the empirical data, but was not useful because it requires specification of three sample statistics (e.g., mean, variance, and skewness) but only two statistics are usually provided.

Finally, like any meta-analysis, Study III is limited by the range of primary studies that have been conducted. Perhaps most importantly, very few intervention studies for MSM have actually measured HIV incidence as an outcome. Thus it was necessary to summarize intervention effects on risk behavior (unprotected sex) rather than on HIV incidence.

8.4 Implications and future research

Study I focuses on correcting the variance of the log odds ratio in group-randomized trials. The correction factors and the traditional *VIF* can be used in power analyses for future studies; researchers planning group-randomized trials should be aware that increasing the number of groups per condition is a much better strategy for improving power than increasing the number of members per group. The correction factors and traditional *VIF* also permit generation of a variance from studies where only one group has been assigned to the intervention condition and one group to comparison.

Additional research can establish the analogous procedures for the log risk ratio, log rate ratio, and standardized mean difference. Further research can also investigate the appropriate correction factors for models that include individual-level and group-level covariates.

Results from Study II suggest that further investigation is warranted using empirical data from other populations besides MSM, and other realms of health and risk behavior besides HIV prevention. What distribution patterns are observed among count variables measuring other risk behaviors, such as number of cigarettes smoked or number of occasions of drug use? What about health promotion behaviors, such as number of

days where subjects engaged in at least 30 minutes of aerobic exercise? Can similar approaches (regression or the method of moments) be used for these outcomes?

The juxtaposition of count and dichotomous measures raises a significant practical question: What measures of behavior are in fact most closely related to risk of transmission of HIV? Epidemiological studies revealed early on that unprotected anal sex and number of partners were related to transmission. But HIV prevention efforts could still benefit from a more refined understanding of transmission dynamics. Newly infected people may be vastly more likely to transmit HIV than those with established (and therefore more immunologically controlled) infections (Jacquez, 1994); would reducing the number of different partners for unprotected sex during the window period of new infection (and high infectiousness) lead to a reduction in the number of transmission events? If so, an optimal measurement of target behavior for risk reduction could be the number of partners for unprotected sex during the specified recall period. To our knowledge, this behavioral message has not yet been tested empirically.

Finally study III shows that behavioral interventions do promote self-reported risk reduction among MSM. A sufficient number of studies have accumulated so that future meta-analyses can be more narrowly targeted, for example by demographic groups (HIV-positive MSM, black MSM, MSM in the developing world), intervention format (small group, individual-level, or community-level), and intervention content (cognitive-behavioral therapy, relapse prevention, community mobilization). Study III revealed a critical lack of effective interventions for African American MSM; our research contributed substantially to a growing awareness of this need, and two interventions for

this population have now been tested, found effective, and recommended for dissemination (Jones 2008, Wilton 2009).

The three studies in concert provide a model and methods for summarizing and comparing the effects of interventions in all fields of health promotion where primary intervention research is being conducted.

9. References

Ajzen I, Fishbein M. *Understanding attitudes and predicting social behavior.* Englewood Cliffs, NJ: Prentice Hall; 1980

Altman DG, Deeks JJ. Odds ratios should be avoided when events are common. *BMJ* 1998;317:1318.

Amirkhanian YA, Kelly JA, Kabakchieva E *et al*. A randomized social network HIV prevention trial with young men who have sex with men in Russia and Bulgaria. *AIDS* 2005;19:1897–1905.

Anand R, Moore J, Jaffe H, *et al.* DNA and protein heterogeneity in serial isolates of human immunodeficiency virus (HIV-1): indication of change in vivo. *Microbios* 1987;52:191-201.

Anderson JE. Condom use and HIV risk among US adults. *Am J Public Health* 2003;93(6):912-914.

Andreasen AR. Social Marketing: Its definition and domain, *Journal of Public Policy & Marketing* 1994;13:108-114.

Anonymous. Interventions can ease dangers of cybersex. Counselors educate while in chatrooms. *AIDS Alert* 2003;18;124-127.

Anonymous. Study confirms the link between MSM and web. San Francisco officials use on-line interventions. *AIDS Alert* 2004;19;19-20.

Atfeld M, Allen TM, Yu XG, *et al.* HIV-1 superinfection despite broad CD8+ T-cell responses containing replication of the primary virus. *Nature* 2002;420:434-439.

Auerbach DM, Darrow WW, Jaffe HW, Curran JW. Cluster of cases of the acquired immune deficiency syndrome. Patients linked by sexual contact. *American Journal of Medicine* 1984;76(3):487-92.

Bandura A. Social cognitive theory and exercise of control over HIV infection. In: DiClemente RJ, Peterson J, eds. Preventing AIDS: theories and methods of behavioral interventions. New York: Plenum; 1994.

Bandura A. *Self-efficacy: the exercise of control*. New York: Freeman; 1997.

Bandura A. Vicarious processes: A case of no-trial learning. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol . 2). New York: Academic; 1965.

Bandura A. *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall; 1977.

Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* 1977;84;191-215.

Bandura A. "Perceived self-efficacy in the exercise of control over AIDS infection." In V.M. Mays, G.W. Albee, & S.S. Schneider (Eds.), *Primary prevention of AIDS: Psychological approaches* (pp. 128-141). Newbury Park, CA:Sage; 1989.

Barré-Sinoussi F, Chermann JC, Rey F, *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 1983;220:868-871.

Beck AT. *Cognitive therapy and the emotional disorders*. New York: International University Press; 1976.

Bellocco R, Pagano M. From the binomial to the Poisson distribution. *Nutrition* 1997;13:842-843.

Binson D, Michaels S, Stall R, Coates TJ, Gagnon JH, Catania JA. Prevalence and social distribution of men who have sex with men: United States and its urban centers. *J Sex Research* 1995;32(3):245-254.

Blake SM, Ledsky R, Goodenow C, Sawyer R, Lohrmann D, Windsor R. Condom availability programs in Massachusetts high schools: relationships with condom use and sexual behavior. *Am J Public Health* 2003;93:955-962.

Bracken MB, Sinclair JC. Avoidable systematic error in estimating treatment effects must not be tolerated. *BMJ* 2003;317:1156.

Brehm JW. A theory of psychological reactance. New York: Academic Press; 1966.

Breslow NE. Extra-Poisson variation in log-linear models. *Appl Statist* 1984;33:38-44.

Breslow NE, Clayton DG. Approximate Inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association* 1993;88;9–25.

Burton A. Enfurvitide approved for defusing HIV. *Lancet Infect Dis* 2003;3:260.

Butera ST, Pisell TL, Limpakarnjanarat K, *et al.* Production of a novel viral suppressive activity associated with resistance to infection among female sex workers exposed to HIV type 1. *AIDS Research and Human Retroviruses* 2001;17:735-744.

Byers RH, Morgan WM, Darrow WW, *et al.* Estimating AIDS infection rates in the San Francisco cohort. *AIDS* 1988;2:207-10.

Campbell MK, Grimshaw JM, Steen IN. Sample size calculations for cluster randomized

trials. *Journal of Health Services Research and Policy* 2000;5:12-16.

Carballo-Diéguez A., Dolezal C, Leu C-S, *et al*. A randomized controlled trial to test an HIV-prevention intervention for Latino gay and bisexual men: Lessons learned. *AIDS Care* 2005;17:314-328.

Catania JA, Morin SF, Canchola J, Pollack L, Chang J, Coates TJ. U.S. priorities—HIV prevention. *Science* 2000;290:717.

CDC. Pneumocystis pneumonia - Los Angeles. *MMWR Morb Mortal Wkly Rep* 1981a; 30:250-252.

CDC. Kaposi's sarcoma and pneumocystis pneumonia among homosexual men - New York City and California. *MMWR Morb Mortal Wkly Rep* 1981b;30:305-308.

CDC. Prevention of acquired immune deficiency syndrome (AIDS):report of inter-agency recommendations. *MMWR Morb Mortal Wkly Rep* 1983;32:101-103.

CDC. Resurgent bacterial sexually transmitted disease in men who have sex with men — King County, Washington, 1997-1999. *MMWR Morb Mortal Wkly Rep* 1999;48:773-7.

CDC. *HIV/AIDS Surveillance Report*, 2001a;13(No.2):tables 9 & 11.

CDC. *HIV/AIDS Surveillance Report*, 2001b;13(No.2):tables 10 & 12.

CDC. HIV incidence among young men who have sex with men — seven US cities, 1994-2000. *MMWR Morb Mortal Wkly Rep* 2001c;50:440-441.

CDC AIDS Community Demonstration Projects (ACDP) Research Group. Community-level HIV intervention in 5 cities: final outcome data from the CDC AIDS Community Demonstration Projects. *Am J Public Health* 1999;89:336-345.

Centers for Disease Control and Prevention. *HIV/AIDS Surveillance Report* 2003;15.

Centers for Disease Control and Prevention. Using the Internet for partner notification of sexually transmitted diseases--Los Angeles County, California, 2003. *MMWR Morbidity & Mortality Weekly Report*, 2004;53:129-31.

Centre for the Evaluation of Health Promotion and Social Interventions. *Review of Effectiveness of Health Promotion Interventions for Men Who Have Sex with Men.* EPI-Centre, Social Science Research Unit, London University Institute of Education; 1996.

Chesney MA, Koblin BA, Barresi PJ, *et al.* An individually tailored intervention for HIV prevention: baseline data from the EXPLORE study. *Am J Public Health*

2003;93(6):933-938.

Chesney MA, Chambers DB, Taylor JM, Johnson LM, Folkman S. Coping effectiveness training for men living with HIV: results from a randomized clinical trial testing a group-based intervention. *Psychosomatic Medicine* 2003;65:1038-1046.

Cheung YB. Zero-inflated models for regression analysis of count data: a study of growth and development. *Statistics in Medicine* 2002;21:1461-1469.

Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine* 2000;19:3127-31.

Choi KH, Coates TJ. Prevention of HIV infection. *AIDS* 1994;8:1371-1389.

Choi KH, Lew S, Vittinghoff E, Catania JA, Barrett DC, Coates TJ. The efficacy of brief group counseling in HIV risk reduction among homosexual Asian and Pacific Islander men. *AIDS* 1996;10:81-87.

Chun TW, Fauci AS. Latent reservoirs of HIV: obstacles to the eradication of virus. *Proc Natl Acad Sci USA* 1999;96:10958-10961.

Ciccarone DH, Kanouse DE, Collins RL, *et al.* Sex without disclosure of positive HIV serostatus in a US probability sample of persons receiving medical care for HIV infection. *Am J Public Health* 2003;93:949-954.

Clavel F, Guetard D, Brun-Vezinet F, *et al.* Isolation of a new human retrovirus from West African patients with AIDS. *Science* 1986;233:343-346.

Cleary PD, Van Devanter N, Steilen M, *et al.* A randomized trial of an education and support program for HIV-infected individuals. *AIDS* 1995;9:1271-1278.

Coates TJ, Aggleton P, Gutzwiller F, *et al.* HIV prevention in developed countries. *Lancet* 1996;348:1143-1148.

Coates TJ, McKusick L, Kuno R., Stites DP. Stress reduction training changed number of sexual partners but not immune function in men with HIV. *American Journal of Public Health* 1989;79:885-7.

Coleman CL. Determinants of HIV and AIDS among young African-American men who have sex with men: a public health perspective. *Journal of National Black Nurses Association* 2003;14:25-29.

Colfax GN, Buchbinder SP, Cornelisse PGA, Vittinghoff E, Mayer K, Celum C. Sexual risk behaviors and implications for secondary HIV transmission during and after HIV seroconversion. *AIDS* 2002;16(11):1529-1535.

Conant M. Losing the war on AIDS. *Journal of the American Academy of Dermatology*, 2004;51(1 Suppl):S47-48.

Cooper H, Hedges LV. The handbook of research synthesis. Russell Sage Foundation, New York; 1994.

Cornfield J. Randomization by group: A formal analysis. *Am J Epidemiol* 1978;108(2):100-102.

Courtenay-Quirk C, Wolitski RJ, Hoff C, Parsons JT. Interests in HIV prevention topics of HIV-seropositive men who have sex with men. *AIDS Education and Prevention* 2003;15:401-412.

Craighead LW, Craighead WE, Kazdin AE, Mahoney MJ. Cognitive and behavioral interventions: An empirical approach to mental health problems. Needham Heights, MA: Allyn & Bacon; 1994.

Crepaz N, Marks G, Mansergh G, Murphy S, Miller LC, Appleby PR. Age-related risk for HIV infection in men who have sex with men: examination of behavioral, relationship, and serostatus variables. *AIDS Educ Prev* 2000;12(5):405-415.

Crosby RA, DiClemente RJ, Wingood GM, Lang D, Harrington KF. Value of consistent condom use: a study of sexually transmitted disease prevention among African American females. *Am J Public Health* 2003;93(6):901-902.

Crosby R, DiClemente RJ, Mettey A. Correlates of recent unprotected anal sex among men having sex with men attending a large sex resort in the South. *Sexually Transmitted Diseases* 2003;30:909-913.

Curran JW. The epidemiology and prevention of the acquired immunodeficiency syndrome. *Annals of Internal Medicine* 1985;103(5):657-62.

Curran JW, Jaffe HW, Hardy AM, *et a*l. Epidemiology of HIV infection and AIDS in the United States. *Science* 1988;239(4840):610-6.

D'Souza MP, Harden VA. Chemokines and HIV-1 second receptors. Confluence of two fields generates optimism in AIDS research. *Nat Med* 1996; 2:1293-1300.

Dalgleish AG, Beverley PCL, Clapham PR, *et al.* The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* 1984;312:763-767.

Daly L. Simple SAS macros for calculation of exact binomial and Poisson confidence limits. *Comput Biol Med* 1992;22:351-361.

Darrow WW, Echenberg DF, Jaffe HW, *et al.* Risk factors for human immunodeficiency virus (HIV) infections in homosexual men. *Am J Public Health* 1987;77(4):479-

483.

Davis KR, Weller SC. The effectiveness of condoms in reducing heterosexual transmission of HIV. *Family Planning Perspectives* 1999;31(6):272-9.

Dean M, Carrington M, Winkler C, *et al.* Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Science* 1996;273:1856-62.

Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine* 2002;21:1575-1600.

Deeks JJ. Odds ratios should be used only in case-control studies and logistic regression analyses. *BMJ* 1998;317:1155-1156.

Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses (chapter 9). In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 (updated September 2008). The Cochrane Collaboration. Available from www.cochrane-handbook.org.

Deeks JJ, on Behalf of the Statistical Methods Working Group of the Cochrane Collaboration. Statistical methods programmed in MetaView. Oxford:Cochrane Library; 1999.

Demmer C. HIV prevention in the era of new treatments. *Health Promotion Practice*, 2003;4:449-456.

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials*, 1986;7:177-88.

Deutsch M, Gerard HB. A study of normative and information social influences on individual judgment. *Journal of Abnormal and Social Psychology*, 1955;51:629-636.

Devare SG, Srinivasan A, Bohan CA, Spira TJ, Curran JW, Kalyanaraman VS. Genomic diversity of the acquired immunodeficiency syndrome retroviruses is reflected in alteration of its translational products. *Proceedings of the National Academy of Sciences of the United States of America* 1986;83(15):5718-22.

Diehr P, Martin DC, Koepsell T, Cheadle A. Breaking the matches in a paired t-test for community interventions when the number of pairs is small. *Statistics in Medicine* 1995;14:1491-1504.

Dilley JW, Woods WJ, Sabatino J, *et al.* Changing sexual behavior among gay male repeat testers for HIV: a randomized, controlled trial of a single-session intervention. *J Acquired Immune Defic Syndr* 2002;30:177-186.

Donald A, Donner A. Adjustments to the Mantel–Haenszel chi-square statistic and odds ratio variance estimator when the data are clustered. *Statistics in Medicine* 1987;6:491-499.

Donner A, Birkett N, Buck C. Randomization by cluster: sample size requirements and analysis. *Am J Epidemiol* 1981;114:906-14.

Donner A. Approaches to sample size estimation in the design of clinical trials — a review. *Statistics in Medicine* 1984;3:199-214.

Donner A. Some aspects of the design and analysis of cluster randomization trials. *Applied Statistics* 1998;47:95-113.

Donner A, Klar N. Methods for comparing event rates in intervention studies when the unit of allocation is a cluster. *American Journal of Epidemiology* 1994;140:279-289.

Donner A, Klar N. Issues in the meta-analysis of cluster randomized trials. *Statistics in Medicine* 2002;21:2971-80.

Donner A, Klar N. Pitfalls of and controversies in cluster randomization trials. *American Journal of Public Health* 2004;94:416-22.

Editorial. What if we had an AIDS vaccine? *Lancet* 2000; 356:85.

Ekstrand ML, Stall RD, Paul JP, *et al.* Gay men report high rates of unprotected anal sex with partners of unknown or discordant HIV status. *AIDS* 1999;13:1525-33.

Elford J, Bolding G, Sherr L. Peer education has no significant impact on HIV risk behaviours among gay men in London. *AIDS* 2001;15:535-538.

Elliott BJ. A *re-examination of the social marketing concept*. Sydney: Elliott & Shanahan Research; 1991.

Emmons CA, Joseph JG, Kessler RC, Wortman CB, Montgomery SB, Ostrow DG. Psychosocial predictors of reported behavior change in homosexual men at risk for AIDS. *Health Education Quarterly* 1986;13:331-345.

EXPLORE Study Team. Effects of a behavioural intervention to reduce acquisition of HIV infection among men who have sex with men: the EXPLORE randomised controlled study. *Lancet* 2004;364:41-50.

Fauci AS. Host factors and the pathogenesis of HIV-induced disease. *Nature* 1996;384:529-534.

Fauci AS. HIV and AIDS: 20 years of science. *Nat Med* 2003;9(7):839-843.

Fee E, Brown TM, Laylor J. One size does not fit all in the transgender community. *Am J Public Health* 2003;93(6):899-900.

Feng Z, Diehr P, Peterson A, McLerran D. Selected statistical issues in group randomized trials. *Annu Rev Public Health* 2001;22:167-87.

Festinger L. A theory of cognitive dissonance. Evanston, IL: Row, Peterson; 1957.

Fihn SD, Boyko EJ, Normand EH, *et al*. Association between spermicide-coated condoms and *Escherichia coli* urinary tract infection in young women. *Am J Epidemiol* 1996;144:512-20.

Fishbein M, Ajzen I. *Belief, attitude, intention, and behavior: An introduction to theory and research.* Boston: Addison-Wesley; 1975.

Fisher DS, Ryan R, Esacove AW, Bishofsky S, Wallis JM, Roffman RA. The social marketing of project ARIES: overcoming challenges in recruiting gay and bisexual males for HIV prevention counseling. *J Homosex* 1996;31(1-2):177-202.

Fisher JD, Fisher WA. The information-motivation-behavioral skills model of AIDS risk behavior change: Empirical support and application. In S. Oskamp and S.C. Thompson (eds.), *Understanding and preventing HIV risk behavior: Safer sex and drug use* (pp. 100-127). Thousand Oaks, CA: Sage; 1996.

Fisher JD, Fisher WA. Changing AIDS-risk behavior. *Psychological Bulletin* 1992;111:455-474.

Fisher JD, Misovich SJ. Social influence and AIDS preventive behavior. In J. Edwards, R.S. Tindall, L. Heath, & E.J. Posavac (Eds.), Social influence processes and prevention (pp. 39-70). New York: Plenum Press; 1990.

Fisher WA. A theory-based framework for intervention and evaluation in STD/HIV prevention. *The Canadian Journal of Human Sexuality* 1997;6(2):151-159.

Flexner C. HIV-protease inhibitors. *N Engl J Med* 1998;338:1281-1292.

Flores-Villanueva PO, Yunis EJ, Delgado JC, *et al.* Control of HIV-1 viremia and protection from AIDS are associated with HLA-Bw4 homozygosity. *Proceedings of the National Academy of Sciences of the United States of America* 2001;98(9):5140-5145.

Flowers P, Hart GJ, Williamson LM, Frankis JS, Der GJ. Does bar-based, peer-led sexual health promotion have a community-level effect amongst gay men in Scotland? *International Journal of STD & AIDS* 2002;13:102-108.

Fontanet AL, Saba J, Chandelying V, *et al.* Protection against sexually transmitted diseases by granting sex workers in Thailand the choice of using the male or female condom: results from a randomized controlled trial. *AIDS* 1998;12:1851-9.

Francis DP, Curran JW, Essex M. Epidemic acquired immune deficiency syndrome: epidemiologic evidence for a transmissible agent. *J Natl Cancer Inst* 1983;71:1-4.

Francis DP, Jaffe HW, Fultz PN, Getchell JP, McDougal JS, Feorino PM. The natural history of infection with the lymphadenopathy-associated virus/human T-lymphotropic virus type III. *Annals of Internal Medicine* 1985;103(5):719-22.

Freire P. *Pedagogy of the oppressed*. New York: Herder & Herder; 1970.

Freudenberg N, Eng E, Flay B, Parcel G. Strengthening individual and community capacity to prevent disease and promote health: In search of relevant theories and principles. *Health Education Quarterly* 1995;22:290-306.

Friedman SR, Ompad DC, Maslow C, *et al.* HIV prevalence, risk behaviors, and high-risk sexual and injection networks among young women injectors who have sex with women. *Am J Public Health* 2003;93(6):902-906.

Gallo RC, Sarin PS, Gelmann EP, *et al*. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science* 1983;220:865-867

Gallus G, Radaelli G, Marchi M. Poisson approximation to a negative binomial process in the surveillance of rare health events. *Methods of Information in Medicine* 1991;30:206-209.

Gao F, Bailes E, Robertson DL, *et al.* Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes. Nature* 1999;397:436-441.

Garber DA, Silvestri G, Feinberg MB. Prospects for an AIDS vaccine: three big questions, no easy answers. *Lancet Infectious Diseases* 2004;4:397-413.

Gardner W, Mulvey EP, Shaw E. Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psych Bulletin* 1995;118:392-404.

Glanz K, Rimer BK. Theory at a glance. Washington:National Institutes of Health; 1995.

Gold RS, Skinner MJ. Gay men's estimates of the likelihood of HIV transmission in sexual behaviours. *Int J STD AIDS* 2001;12(4):245-255.

Gold RS, Rosenthal DA. Preventing unprotected anal intercourse in gay men: A comparison of two intervention techniques. *International Journal of STD & AIDS*

1995;6:89-94.

Gold RS, Rosenthal DA. Examining self-justifications for unsafe sex as a technique of AIDS education: the importance of personal relevance. *International Journal of STD & AIDS* 1998;9:208-213.

Golombok S, Harding R, Sheldon J. An evaluation of a thicker versus a standard condom with gay men. *AIDS* 2001;15(2):245-250.

Gottlieb MS, Schroff R, Schanker HM, *et al.* Pneumocystis carinii pneumonia and mucosal candidiasis in previously health homosexual men: evidence of a new acquired cellular immunodeficiency. *N Engl J Med* 1981;305:1425-1431.

Goulder PJ, Walker BD. HIV-1 superinfection - a word of caution. *N Engl J Med* 2002;347:756-758.

Gray RH, Li X, Wawer MJ, *et al.* Stochastic simulation of the impact of antiretroviral therapy and HIV vaccines on HIV transmission; Rakai, Uganda. *AIDS* 2003;17, 1941-51.

Greenland S. Meta-analysis. pp. 643-673 In: Rothman KJ, Greenland S eds. *Modern Epidemiology*. Philadelphia: Lippincott Williams and Wilkins; 1998.

Gross M. When plagues don't end. *Am J Public Health* 2003a;93(6):861-862.

Gross M. The second wave will drown us. *Am J Public Health* 2003b;93(6):872-881.

Grosskurth H, Mosha F, Todd J, *et al.* Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomised controlled trial. *Lancet* 1995;346:530-536.

Guenther-Grey CA, Varnell S, Weiser JI, *et al.* Trends in Sexual Risk-Taking among Urban Young Men Who Have Sex with Men, 1999-2002. *Journal of the National Medical Association* 2005;97(7):38S–43S.

Gut A. *Probability: A Graduate Course*. Springer-Verlag; 2005.

Hall HI, Song R, Rhodes P, *et al*. Estimation of HIV incidence in the United States. *JAMA* 2008;300:520-29.

Hanenberg R, Rojanapithayakorn W. Changes in prostitution and the AIDS epidemic in Thailand. *AIDS Care* 1998;10:69-79.

Hannan PJ, Murray DM. Gauss or Bernoulli? A Monte Carlo comparison of the performance of the linear mixed model and the logistic mixed model analyses in

simulated community trials with a dichotomous outcome variable at the individual level. *Evaluation Review* 1996;20 (3):338-352.

Harding R, Bensley J, Corrigan N, *et al*. Outcomes and lessons from a pilot RCT of a community-based HIV prevention multi-session group intervention for gay men. *AIDS Care* 2004;16:581-585.

Hardy AM, Allen JR, Morgan WM, Curran JW. The incidence rate of acquired immunodeficiency syndrome in selected populations. *JAMA* 1985;253(2):215-20.

Hardy AM, Rauch K, Echenberg D, Morgan WM, Curran JW. The economic impact of the first 10,000 cases of acquired immunodeficiency syndrome in the United States. *JAMA* 1986;255(2):209-11.

Hart GJ. The evaluation of behavioural interventions for gay men: obstacles to evidence-based prevention. *International Journal of STD & AIDS* 1996;7(Supplement 2):25-29.

Hasselblad V, Hedges LV. Meta-analysis of screening and diagnostic tests. *Psych Bulletin* 1995;117:167—78.

Hays RB, Peterson JL. HIV Prevention for Gay and Bisexual men in Metropolitan Cities. Pages 267-296 In *Preventing AIDS: Theories and Methods of Behavioral Interventions*, Ralph J. DiClemente and John L. Peterson (eds.). New York: Plenum Press; 1994.

Hedges LV. Fixed effects models. pp. 285-299 In:Cooper H, Hedges LV eds. *The handbook of research synthesis*. New York: Russell Sage Foundation; 1994.

Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychological Methods* 1998;3:486-504.

Herbst JH, Sherba RT, Crepaz N, *et al*. A meta-analytic review of HIV behavioral interventions for reducing sexual risk behavior of men who have sex with men. *J Acquir Immune Defic Syndr* 2005;39, 228-241.

Higgins DL, O'Reilly K, Tashima N *et al*. Using formative research to lay the foundation for community level HIV prevention efforts: an example from the AIDS Community Demonstration Projects. *Public Health Rep* 1996;111(suppl1):28-35.

Hochbaum G, Sorenson J, Lorig K. Theory in health education practice. *Health Education Quarterly* 1992;19(3):295-313.

Hoff CC, Faigeles B, Wolitski RJ, *et al*. Sexual risk of HIV transmission is missed by traditional methods of data collection. *AIDS* 2004;18:340-342.

Hoff CC, Kegeles SM, Acree M, *et al*. Looking for men in all the wrong places...: HIV prevention small-group programs do not reach high risk gay men. *AIDS* 1997;11:829-30.

Holtgrave DR, Qualls NL, Curran JW, *et al*. An overview of the effectiveness and efficiency of HIV prevention programs. *Public Health Reports* 1995;110(2):134-146.

Hsieh FY. Sample size formulae for intervention studies with the cluster as unit of randomization. [erratum appears in Stat Med 1997; 16:1300] *Statistics in Medicine* 1988;7:1195-1201.

Hsieh FY, Lavori PW, Cohen HJ, Feussner JR. An Overview of Variance Inflation Factors for Sample-Size Calculation. *Evaluation & the Health Professions* 2003;26:239-257.

Hughes JN, Kemenoff SN. The cognitive-behavioral approach. In R .C. D'Amato and B. A. Rothlisberg [Eds.] *The Quest for Answers: A comparative study of intervention models through case study* (pp. 269-303) White Plains, NY: Longman Publishing Group; 1992

Imrie J, Stephenson JM, Cowan FM, *et al.* A cognitive behavioural intervention to reduce sexually transmitted infections among gay men: randomised trial. *BMJ* 2001;322:1451-6.

Jacquez JA, Koopman JS, Simon CP, Longini IM. Role of the primary infection in epidemics of HIV infection in gay cohorts. *J Acquir Immune Defic Syndr* 1994;7:1169-1184.

Jaffe HW, Choi K, Thomas PA, *et al.* National case-control study of Kaposi's sarcoma and Pneumocystis carinii pneumonia in homosexual men: Part 1. Epidemiologic results. *Annals of Internal Medicine* 198399(2):145-51.

Jaffe HW, Francis DP, McLane MF, *et al.* Transfusion-associated AIDS: serologic evidence of human T-cell leukemia virus infection of donors. *Science* 1984;223(4642):1309-12.

Jaffe HW, Hardy AM, Morgan WM, Darrow WW. The acquired immunodeficiency syndrome in gay men. *Annals of Internal Medicine* 1985;103(5):662-4.

Jaffe HW, Sarngadharan MG, DeVico AL, *et al.* Infection with HTLV-III/LAV and transfusion-associated acquired immunodeficiency syndrome. Serologic evidence of an association. *JAMA* 1985;254(6):770-3.

Janz NK, Becker MH. The health belief model: a decade later. *Health Education Quarterly* 1984;11:1-47.

Jayaraman GC, Read RR, Singh A. Characteristics of individuals with male-to-male and heterosexually acquired infectious syphilis during an outbreak in Calgary, Alberta, Canada. *Sexually Transmitted Diseases* 2003;30(4):315-319.

Johnson RE, Lawrence DN, Evatt BL, *et al.* Acquired immunodeficiency syndrome among patients attending hemophilia treatment centers and mortality experience of hemophiliacs in the United States. *Am J Epidemiology* 1985;121(6):797-810.

Johnson WD, Hedges LV, Diaz RM. Interventions to modify sexual risk behaviors for preventing HIV infection in men who have sex with men (Cochrane Review). In:The Cochrane Library, Issue 3, Chichester, UK:John Wiley & Sons, Ltd.; 2003.

Johnson WD, Hedges LV, Ramirez G, *et al.* HIV prevention research for men who have sex with men: a systematic review and meta-analysis. *J Acquir Immune Defic Syndr* 2002a;30(suppl 1):S118-S129.

Johnson WD, Holtgrave DR, McClellan WM, Flanders WD, Hill AN, Goodman M. HIV intervention research for men who have sex with men: a 7-year update. *AIDS Education and Prevention* 2005;17(6):568–589.

Johnson WD, Semaan S, Hedges LV, Ramirez G, Mullen PD, Sogolow E. A protocol for the analytical aspects of a systematic review of HIV prevention research. *J Acquir Immune Defic Syndr* 2002b;30(suppl 1):S62-72.

Jones KT, Gray P, Whiteside YO, Wang T, Bost D, Dunbar E, Foust E, Johnson WD. Evaluation of an HIV prevention intervention adapted for Black men who have sex with men. *Am J Public Health* 2008;98:1043-50.

Joseph JG, Montgomery SB, Emmons C, *et al*. Magnitude and determinants of behavioral risk reduction: longitudinal analysis of a cohort at risk for AIDS. *Psychol Health* 1987;1:73-95.

Kalichman SC, Carey MP, Johnson BT. Prevention of sexually transmitted HIV infection: A meta-analytic review of the behavioral outcome literature. *Ann Behav Med* 11996;8(1):5-15.

Kalichman SC, Rompa D, Cage M, *et al.* Effectiveness of an intervention to reduce HIV transmission risks in HIV-positive people. *Am J Prev Med* 2001;21:84-92.

Katz KA. The (relative) risks of using odds ratios. *Arch Dermatol* 2006;142:761-764.

Kaul R, Rowland-Jones SL, Kimani J, *et al.* Late seroconversion in HIV-resistant Nairobi prostitutes despite pre-existing HIV-specific CD8+ responses. *J Clin Invest* 2001;107:341-349.

Kegeles SM, Hart GJ. Recent HIV-prevention interventions for gay men: individual, small-group and community-based studies. *AIDS* 1998;12 (supplement A):S209-S215.

Kegeles SM, Hays RB, Coates TJ. The Mpowerment Project :a community-level HIV prevention intervention for young gay men. *Am J Public Health* 1996;86:1129-1136.

Kegeles SM, Rebchook GM, Hays RB, Pollack LM. Staving off increases in young gay/bisexual men's risk behavior in the HAART era. In: XIV International Conference on AIDS. Barcelona, (Abstract MoPeD3549; 2002).

Kelly JA, Murphy DA, Bahr R, *et al.* Outcome of cognitive-behavioral and support group brief therapies for depressed, HIV-infected persons. *Am J Psychiatry* 1993;150:1679-1686.

Kelly JA, Murphy DA, Sikkema KJ, *et al.* Randomised, controlled, community-level HIV-prevention intervention for sexual-risk behaviour among homosexual men in US cities. *Lancet* 1997;350:1500-1505.

Kelly JA, St. Lawrence JS, Diaz YE, *et al.* HIV risk behavior reduction following intervention with key opinion leaders of population: an experimental analysis. *Am J Public Health* 1991;8:168-171.

Kelly JA, St. Lawrence JS, Hood HV, Brasfield TL. Behavioral intervention to reduce AIDS risk activities. *Journal of Consulting and Clinical Psychology* 1989;57:60-67.

Kerry SM. Unequal cluster sizes in UK general practice: implications for sample size calculations. *Statistics in Medicine* 2001;20:377-390.

Kilmarx PH, Limpakarnjanarat K, Mastro TD, *et al.* HIV-1 seroconversion in a prospective study of female sex workers in northern Thailand: continued high incidence among brothel-based women. *AIDS* 1998;12:1889-98.

Kingsley LA, Detels R, Kaslow R, *et al.* Risk factors for seroconversion to human immunodeficiency virus among male homosexuals: Results from the Multicenter AIDS Cohort Study. *Lancet* 1987;329:345-9.

Kinsey AC, Pomeroy WE, Martin CE. *Sexual behavior in the human male*. Philadelphia: W.B. Saunders; 1948.

Kinter A, Arthos J, Cicala C, Fauci AS. Chemokines, cytokines and HIV: a complex network of interactions that influence HIV pathogenesis. *Immunol Rev* 2000;177:88-98.

Kish L. *Survey Sampling*. New York: John Wiley and Sons; 1965.

Klar N, Donner A. Current and future challenges in the design and analysis of cluster randomization trials. *Statistics in Medicine* 2001;20:3729-40.

Klar N, Donner A. The merits of matching in community intervention trials:a cautionary tale. *Statistics in Medicine* 1997;16:1753-64.

Klatzmann D, Champagne E, Chamaret S, *et al.* T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* 1984;312:767-768.

Koblin BA, Chesney MA, Husnik MJ, *et al.* High-risk behaviors among men who have sex with men in 6 US cities: baseline data from the EXPLORE study. *Am J Public Health* 2003;93(6):926-932.

Koopman J, Jacquez J, Welch G *et al.* The role of early HIV infection in the spread of HIV through populations. *J Acquir Immune Defic Syndr* 1997;14:249-258.

Korber B, Gaschen B, Yusim K, *et al.* Evolutionary and immunological implications of contemporary HIV-1 variation. *Brit Med Bull* 2001;58:19-42.

Korenromp EL, Bakker R, de Vlas SJ, *et al.* HIV dynamics and behaviour change as determinants of the impact of sexually transmitted disease treatment on HIV transmission in the context of the Rakai trial. *AIDS* 2002;16:2209-2218.

Kotler P, Roberto E. Social marketing: Strategies for changing public behavior. New York: Free Press; 1989.

Kotler P, Zaltman G. Social marketing: An approach to planned social change. *Journal of Marketing* 1971;35:3-12.

Kretzschmar M, Adler F. Aggregated distributions in models for patchy populations. *Theoretical Population Biology* 1993;43:1-30.

Leone P, Hightow L, Foust E, *et al.* HIV transmission among black college student and non-student men who have sex with men --- North Carolina, 2003. *MMWR Morbid Mortal Wkly Report* 2004;53:731-734.

Levy JA, Hoffman AD, Kramer SM, *et al.* Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* 1984;225:840-842.

Lewsey JD, Gilthorpe MS, Bulman JS, Bedi R. Is modeling dental caries a "normal" thing to do? *Community Dental Health* 2000;17:212-217.

Lister NA, Smith A, Tabrizi S, *et al* . Screening for Neisseria gonorrhoeae and

Chlamydia trachomatis in men who have sex with men at male-only saunas. *Sexually Transmitted Diseases* 2003;30:886-889.

Lui KJ, Darrow WW, Rutherford GW. A model-based estimate of the mean incubation period for AIDS in homosexual men. *Science* 1988;240:1333-1335.

Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine* 2001;20:641-654.

Maiman LA, Becker MH. The health belief model: Origins and correlates in psychological theory. *Health Education Monographs* 1974;2:336-353.

Malebranche DJ. Black men who have sex with men and the HIV epidemic: next steps for public health. *Am J Public Health* 93(6):2003;862-864.

Mann JM, Quinn TC, Francis H, *et al.* Prevalence of HTLV-III/LAV in household contacts of patients with confirmed AIDS and controls in Kinshasa, Zaire. *JAMA* 1986;256(6):721-4.

Maris E. Covariance adjustment versus gain scores revisited. *Psychol Methods* 1998;3:309–27.

Marlatt GA, Gordon JR. Determinants of relapse: Implications for the maintenance of behavior change, in P. O. Davidson and S. M. Davidson (eds.), *Behavioral Medicine: Changing Health Lifestyle*. New York: Brunner/ Mazel; 1980.

Marlatt GA, Gordon JR (Eds.). *Relapse prevention: Maintenance strategies in the treatment of addictive behaviors*. New York: Guilford Press; 1985.

Martin DC, Diehr P, Perrin EB, Koepsell TD. The effect of matching on the power of randomized community intervention studies. *Statistics in Medicine* 1993;12:329-338.

Masur H, Michelis MA, Greene JB, *et al.* An outbreak of community-acquired pneumocystis carinii pneumonia: initial manifestations of cellular dysfunction. *N Engl J Med* 1981;305:1431-1438.

McFarland W, Busch MP, Kellogg TA, *et al.* Detection of early HIV infection and estimation of incidence using a sensitive/less-sensitive enzyme immunoassay testing strategy at anonymous counseling and testing sites in San Francisco. *J Acquir Immune Defic Syndr* 1999;22:484–9.

McGuire WJ. The nature of attitudes and attitude change. In G. Lindzey & E Aronson (eds) Handbook of social psychology (vol 3) Reading, MA:Addison-Wesley; 1969.

McMichael AJ, Hanke T. HIV vaccines 1983-2003. *Nat Med* 2003;9(7):874-880.

Merton RK. The unanticipated consequences of purposive social action. *American Sociological Review* 1936;1:894-904.

Miaou SP. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis and Prevention* 1994;26:471-482.

Middlestadt S, Hoffman C, D'Andrea EM, Ruscavage D. *What Intervention Studies Say About Effectiveness: A Resource for HIV Prevention Community Planning Groups.* Academy for Educational Development, Washington, DC; 1996.

Miller B, Stansfield SK, Zack MM, *et al.* The syndrome of unexplained generalized lymphadenopathy in young men in New York City. Is it related to the acquired immune deficiency syndrome? *JAMA* 1984;251:242-6.

Miller RL, Klotz D, Eckholdt HM. HIV prevention with male prostitutes and patrons of hustler bars: replication of an HIV preventive intervention. *American Journal of Community Psychology* 1998;26:97-131.

Miller WR, Benefield GS, Tonigan GS. Enhancing motivation for change in problem drinking: A controlled comparison of two therapist styles. *Journal of Consulting and Clinical Psychology* 1993;61:455-461.

Miller WR, Sovereign RG. The check-up: A model for early intervention in addictive behaviors. In T. Loberg, W.R. Miller, P.E. Nathan, and G. A. Marlatt (Eds.), *Addictive behaviors: Prevention and early intervention* (pp. 219-311). Amsterdam: Swets and Zeitlinger; 1989.

Miller WR, Sovereign RG, Krege B. Motivational interviewing with problem drinkers: II. The Drinker's Check-up as a preventive intervention. *Behavioral Psychotherapy* 1988;16:251-268.

Miller WR, Rollnick W. *Motivational interviewing: Preparing people to change addictive behavior.* New York: Guilford; 1991.

Millett G. More effort needed with 'down low' males. CDC expert discusses issue in Q & A. *AIDS Alert* 2004;19:42-43.

Minkler M. Improving health through community organization, in Glanz, K., Lewis, F.M. & Rimer, B.K. (Eds.) *Health behavior and health education: Theory, research, and practice* (pp 257-287). San Francisco: Jossey-Bass Publishers; 1990.

Montgomery SB, Joseph JG, Becker MH, *et al*. The Health Belief Model in understanding compliance with preventive recommendations AIDS: how useful?

AIDS Educ Prev 1989;1(4):303-23.

Moore H, Summerbell C, Vail A, Greenwood DC, Adamson AJ. The design features and practicalities of conducting a pragmatic cluster randomized trial of obesity management in primary care. *Statistics in Medicine* 2001;20:331-40.

Murray DM. *Design and Analysis of Group-Randomized Trials.* New York: Oxford University Press; 1998.

Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. *Applied Linear Statistical Models.* Chicago: Irwin; 1996.

NIAID. *Workshop summary: scientific evidence on condom effectiveness for sexually transmitted disease prevention, June 12-13, 2000.* Hyatt Dulles Airport, Herndon, Virginia; 2001.

Nicholson JK, McDougal JS, Jaffe HW, *et al.* Exposure to human T-lymphotropic virus type III/lymphadenopathy-associated virus and immunologic abnormalities in asymptomatic homosexual men. *Ann Int Med* 1985;103(1):37-42.

O'Brien SJ, Moore JP. The effect of genetic variation in chemokines and their receptors on HIV transmission and progression to AIDS. *Immunol Rev* 2000;177:99-111.

O'Leary A, DiClemente RJ, Aral SO. Reflections on the design and reporting of STD/HIV behavioral intervention research. *AIDS Educ Prevent* 1997;9(suppl A):1-14.

Oakley A, Fullerton D, Holland J. Behavioural interventions for HIV/AIDS prevention. *AIDS* 1995;9:479-486.

Oehlert GW. A note on the delta method. *The American Statistician* 1992;46(1):27-29.

Palacios-Jimenez L, Shernoff M. *Facilitator's Guide to Eroticizing safer sex.* Gay Men's Health Crisis, New York; 1986.

Pals SL, Murray DM, Alfano CM, Shadish WR, Hannan PJ, Baker WL. Individually randomized group treatment trials: a critical appraisal of frequently used design and analytic approaches. *American Journal of Public Health.* 2008;98:1418-24.

Patterson TL, Shaw WS, Semple SJ. Reducing the sexual risk behaviors of HIV+ individuals: outcome of a randomized controlled trial. *Ann Behav Med* 2003;25:137-145.

Paul JP, Catania J, Pollack L, *et al.* Suicide attempts among gay and bisexual men: lifetime prevalence and antecedents. *Am J Public Health* 2002;92:1338-1345.

Paxton WA, Liu R, Kang S, *et al.* Reduced HIV-1 infectability of CD4+ lymphocytes from exposed-uninfected individuals: Association with low expression of CCD5 and high production of beta-chemokines. *Virology* 1998;244:66-73.

Perry S, Fishman B, Jacobsberg L, Young J, Frances A. Effectiveness of psychoeducational interventions in reducing emotional distress after Human Immunodeficiency Virus antibody testing. *Arch Gen Psychiatry* 1991;48:143-147.

Peterman TA, Curran JW. Sexual transmission of human immunodeficiency virus. *JAMA* 1986;256(16):2222-6.

Peterman TA, Drotman DP, Curran JW. Epidemiology of the acquired immunodeficiency syndrome (AIDS). *Epidemiologic Reviews* 1985;7:1-21.

Peterman TA, Jaffe HW, Feorino PM, *et al.* Transfusion-associated acquired immunodeficiency syndrome in the United States. *JAMA* 1985;254(20):2913-7.

Peterman TA, Stoneburner RL, Allen JR, Jaffe HW, Curran JW. Risk of human immunodeficiency virus transmission from heterosexual adults with transfusion-associated infections. [erratum appears in J*AMA 1989*; 262(4):502]. *JAMA* 1988;259(1):55-8.

Peterson JL, Ostrow DG, McKirnan DJ. Behavioral interventions for the primary prevention of HIV infection among homosexual and bisexual men. *Journal of Primary Prevention* 1991;12;1:19-34.

Peterson JL, Coates TJ, Catania J, *et al.* Evaluation of an HIV risk reduction intervention among African-American homosexual and bisexual men. *AIDS* 1996;10:319-325.

Picciano JF, Roffman RA, Kalichman SC, Rutledge SE, Berghuis JP. A telephone based brief intervention using motivational enhancement to facilitate HIV risk reduction among MSM: A pilot study. *AIDS & Behavior* 2001;5:251-262.

Pinkerton SD, Abramson PR. Effectiveness of condoms in preventing HIV transmission. *Soc Sci Med* 1997;44(9):1303-12.

Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: Applications to addictive behaviors. *American Psychologist* 1992;47:1102-1114.

Prochaska JO, Redding CA, Harlow LL, Rossi JS, and Velicer, WF. The transtheoretical model of change and HIV prevention: A review. *Health Education Quarterly* 1994;21:471-486.

Puckett SB, Bye LL. *The Stop AIDS Project: an interpersonal AIDS-prevention program.* San Francisco: the Stop AIDS Project, Inc.; 1987.

Quadland MC, Shattis W, Schuman R, Jacobs R, D'Eramo J. *The 800 Men Study: systematic evaluation of AIDS prevention programs*. New York: Gay Men's Health Crisis Inc.; 1988.

Rabkin CS, Thomas PA, Jaffe HW, Schultz S. Prevalence of antibody to HTLV-III/LAV in a population attending a sexually transmitted diseases clinic. *Sexually Transmitted Diseases* 1987;14:48-51.

Richardson JL, Milam J, McCutchan A, *et al.* Effect of brief safer-sex counseling by medical providers to HIV-1 seropositive patients: a multi-clinic assessment. *AIDS* 2004;18:1179-1186.

Robins AG, Dew MA, Kingsley LA, Becker JT. Do homosexual and bisexual men who place others at potential risk for HIV have unique psychological profiles? *AIDS Educ Prev* 1997;9:239-51.

Roffman RA, Picciano JF, Ryan R, *et al.* HIV-prevention group counseling delivered by telephone: an efficacy trial with gay and bisexual men. *AIDS and Behavior* 1997;1:137-154.

Roffman RA, Stephens RS, Curtin L, *et al.* Relapse prevention as an interventive model for HIV risk reduction in gay and bisexual men. *AIDS Education and Prevention* 1998;10:1-18.

Rogers E. Diffusion theory: A theoretical approach to promote community-level change. (pp. 57-65) In J.L. Peterson & R.J. DiClemente (eds.), *Handbook of HIV prevention*. New York: Kluwer Academic / Plenum; 2000.

Rogers EM. *Diffusion of innovations*, 1st edition. New York: Free Press; 1962.

Rogers EM. *Diffusion of innovations*, 3rd edition. New York: Free Press; 1983.

Rojanapithayakorn W, Hanenberg R. The 100% condom program in Thailand. *AIDS* 1996;10:1-7.

Rose CE, Martin SW, Wannemuehler KA, Plikaytis BD. On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. *Journal of Biopharmaceutical Statistics* 2006;16(4):463-81.

Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the health belief model. *Health Education Quarterly* 1988;15:175-183.

Rosenstock IM, Strecher VJ, Becker MH. The health belief model and HIV risk behavior change. In R.J. DiClemente and J. L. Peterson (eds.), *Preventing AIDS: Theories and methods of behavioral interventions* (pp. 5-24). New York: Plenum; 1994.

Rosenthal R, Rosnow RL. *Essentials of Behavioral Research: Methods and Data Analysis*. Boston:McGraw Hill; 1991.

Ross MW, Rosser BRS. Education and AIDS risks: a review. *Health Education Research* 1989;4:273-284.

Rosser BRS. Evaluation of the efficacy of AIDS education interventions for homosexually active men. *Health Education Research* 1990;5:299-308.

Rosser BRS, Bockting WO, Rugg DL, *et al.* A randomized controlled intervention trial of a sexual health approach to long-term HIV risk reduction in men who have sex with men: effects of the intervention on unsafe sexual behavior. *AIDS Education & Prevention* 2002;14(Suppl3):59-71.

Rotheram-Borus MJ, Lee MB, Murphy DA, *et al*. Efficacy of a preventive intervention for youths living with HIV. *Am J Public Health* 2001;91:400-405.

Rotheram-Borus MJ, Swendeman D, Comulada WS, Weiss RE, Lee M, Lightfoot M. Prevention for substance-using HIV-positive young people: telephone and in-person delivery. *J Acquir Immune Defic Syndr* 2004;37:S68-S77.

Rugpao S, Beyrer C, Tovanabutra S, *et al.* Multiple condom use and decreased condom breakage and slippage in Thailand. *J Acquir Immune Defic Syndr and Human Retrovirology* 1997;14:169-73.

Sampaio M, Brites C, Stall R, Hudes ES, Hearst N. Reducing AIDS risk among men who have sex with men in Salvador, Brazil. *AIDS & Behavior* 2002;6:173-181.

Sarngadharan MG, Popovic M, Bruch L, Schupbach J, Gallo RC. Antibodies reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS. *Science* 1984;224:506-508.

Scarlatti G, Tresoldi E, Bjorndal A, *et al. In vivo* evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression. *Nat Med* 1997;3:1259-1265.

Seligman MEP. *Helplessness: On depression, development, and death*. San Francisco: Freeman; 1975.

Semaan S, Kay L, Strouse D, Sogolow E, Mullen PD, Neumann MS, *et al.* A profile of U.S.-based trials of behavioral and social interventions for HIV risk reduction. *J Acquir Immune Defic Syndr* 2002;30(suppl 1):S30-S50.

Semple SJ, Patterson TL, Grant I. Binge use of methamphetamine among HIV-positive men who have sex with men: pilot data and HIV prevention implications. *AIDS Education & Prevention* 2003;15:133-47.

Senn SJ. Covariance analysis in generalized linear measurement error models. *Statistics in Medicine* 1990;9:583–6.

Severini TA. *Elements of Distribution Theory* (example 13.3). Cambridge University Press; 2005.

Shepherd J, Weare K, Turner G. Peer-led sexual health promotion with young gay and bisexual men: results of The HAPEER Project. *Health Education*, 1997;97:204-212.

Shoptaw S, Reback CJ, Peck JA, *et al*. Behavioral treatment approaches for methamphetamine dependence and HIV-related sexual risk behaviors among urban gay and bisexual men. *Drug and Alcohol Dependence* 2005;78:125-134.

Silenzio VMB. Lesbian, gay, and bisexual health in cross-cultural perspective. *J Gay Lesbian Med Assoc* 1997;1:75-86.

Silenzio VMB. Anthropological assessment for culturally appropriate interventions targeting men who have sex with men. *Am J Public Health* 2003;93:867-871.

Smith MW, Dean M, Carrington M *et al.* Contrasting genetic influence of CCR2 and CCR5 variants on HIV infection and disease progression. *Science* 1997;277:959-965.

Sorensen JL, Dilley J, London J, Okin RL, Delucchi KL, Phibbs CS. Case management for substance abusers with HIV/AIDS: a randomized clinical trial. *American Journal of Drug and Alcohol Abuse* 2003;29:133-150.

Sormani MP, Bruzzi P, Miller DH, Gasperini C, Barkhof F, Filippi M. Modelling MRI enhancing lesion counts in multiple sclerosis using a negative binomial model: implications for clinical trials. *Journal of the Neurological Sciences* 1999;163:74-80.

Srinivasan A, York D, Ranganathan P, *et al.* Transfusion-associated AIDS: donor-recipient human immunodeficiency virus exhibits genetic heterogeneity. *Blood* 1987;69(6):1766-70.

Sriwanantha B, Hodges T, Mastro TD, *et al.* HIV-specific cytotoxic T lymphocytes, HLA-A11 and chemokine related factors may act synergistically to determine HIV resistance in CCR5 D32 negative female sex workers in Chiang Rai, northern Thailand. *AIDS Res Hum Retroviruses* 2001;17:719-734.

Stall R, Mills TC, Williamson J *et al.* Association of co-occurring psychosocial health problems and increased vulnerability to HIV/AIDS among urban men who have sex with men. *Am J Public Health* 2003;93(6):939-942.

Stall RD, Coates TJ, Hoff C. Behavioral risk reduction for HIV infection among gay and bisexual men. *American Psychologist* 1988;43:878-885.

Stall RD, Paul JP, Barrett DC, Crosby GM, Bein E. An outcome evaluation to measure changes in sexual risk-taking among gay men undergoing substance use disorder treatment. *Journal of Studies on Alcohol* 1999;60:837-845.

Stokes JP, Vanable P, McKirnan DJ. Comparing gay and bisexual men on sexual behavior, condom use, and psychosocial variables related to HIV/AIDS. *Arch Sex Behav* 1997;26(4):383-397.

Stott J, Hahn BH. Vaccines and immunology. *AIDS* 1999;13(suppl A):S103-S104.

Suissa S. Binary methods for continuous outcomes :a parametric alternative. *J Clin Epidemiol* 1991;44:241–248.

Thompson B, Pertschuck M. Community intervention and advocacy, In Ockene, J.K. & Ockene, J.S. (Eds.), *Prevention of coronary heart disease* (pp. 493-515). Boston:Little, Brown; 1992.

Tudiver F, Myers T, Kurtz RG, *et al.* The Talking Sex Project: results of a randomized controlled trial of small-group AIDS education for 612 gay and bisexual men. *Evaluation & the Health Professions* 1992;15:26-42.

Valdiserri RO, Ogden L, Hansswen R, Onorato I. Aligning the U.S. HIV prevention budget with national goals to reduce new infections by half. *XIV International AIDS Conference* Barcelona 2002, abstract WeOrG1369.

Valdiserri RO, Lyter DW, Leviton LC, Callahan CM, Kingsley LA, Rinaldo CR. AIDS prevention in homosexual and bisexual men: results of a randomized trial evaluating two risk reduction interventions. *AIDS* 1989;3:21-26.

Valleroy LA, MacKellar DA, Karon JM *et al.* HIV Prevalence and Associated Risks in Young Men Who Have Sex With Men. *JAMA* 2000;284(2):198-204.

Van Damme L, Ramjee G, Vuylsteke B, *et al.* Effectiveness of COL-1492, a nonoxynol-9 vaginal gel, on HIV-1 transmission in female sex workers: a randomised controlled trial. *Lancet* 2002;360(9338):971-77. Correction appears in *Lancet* 2002; 360(9348)

Van Breukelen GJP. ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology* 2006;59:920–925.

Varnell S, Murray DM, Janega JB, Blitstein JL. Design and analysis of group-randomized trials:a review of recent practices. *Am J Public Health* 2004;94:393-

399.

Vittinghoff E, Douglas J, Judson F, McKirnan D, MacQueen K, Buchbinder SP. Per-contact risk of Human Immunodeficiency Virus transmission between male sexual partners. *Am J Epidemiol* 1999;150(3):306-311.

Vittinghoff E, Hessol NA, Bacchetti P, Fusaro RE, Holmberg SD, Buchbinder SP. Cofactors for HIV disease progression in a cohort of homosexual and bisexual men. *J Acquir Immune Defic Syndr* 2001;27(3):308-314.

Wackerly DD, Mendenhall W, Scheaffer RL. *Mathematical Statistics with Applications*. 5th edn. Belmont, California: Duxbury Press, Wadsworth Publishing Company; 1996.

Wasserheit JN. Epidemiological synergy. Interrelationships between human immunodeficiency virus infection and other sexually transmitted diseases. *Sex Transm Dis* 1992;19:61-77.

Wawer MJ, Sewankambo NK, Serwadda D, *et al.* Control of sexually transmitted diseases for AIDS prevention in Uganda: a randomised community trial. *Lancet* 1999;353:525-535.

Weinhart LS, Carey MP, Johnson BT, Bichham N. Effects of HIV counseling and testing on sexual risk behavior. *Am J Public Health* 1999;89:1397-1405.

Weinstein ND. Testing four competing theories of health-protective behavior. *Health Psychology* 1993;12:324-333.

Whitehead A, Baily AJ, Elbourne D. Combining summaries of binary outcomes with those of continuous outcomes in a meta-analysis. *Journal of Biopharmaceutical Statistics* 1999;9:1-16.

Wilton L, Herbst JH, Coury-Doniger P, *et al*. Efficacy of an HIV/STI Prevention Intervention for Black Men Who Have Sex With Men:Findings from the *Many Men, Many Voices* (*3MV*) Project. *AIDS and Behavior* 2009;13:532-544.

Wolfinger R, O'Connell M. Generalized Linear Mixed Models:A Pseudo-Likelihood Approach. *Journal of Statistical Computation and Simulation* 1993;4:233-243.

Wolitski RJ, Valdiserri RO, Denning PH, Levine WC. Are we headed for a resurgence of the HIV epidemic among men who have sex with men? *Am J Public Health* 2001;91:883-888.

Wolitski RJ, Gómez CA, Parsons JT, and the SUMIT Study Group. Effects of a peer-led behavioral intervention to reduce HIV transmission and promote serostatus disclosure among HIV-seropositive gay and bisexual men. *AIDS* 2005;19(Suppl

1):S99-S109.

Wong ML, Chan RKW. A prospective study of pharyngeal gonorrhoea and inconsistent condom use for oral sex among female brothel-based sex workers in Singapore. *International Journal of STD & AIDS* 1999;10(9):595-9.

Wulfert E, Wan CK. Safer sex intentions and condom use viewed from a health belief, reasoned action, and social cognitive perspective. *Journal of Sex Research* 1995;32:299-311.

Xiridou M, Geskus R, de Wit J, Coutinho R, Kretzschmar M. The contribution of steady and casual partnerships to the incidence of HIV infection among homosexual men in Amsterdam. *AIDS* 2003;17(7):1029-1038.

Zea MC, Reisen CA, Diaz RM. Methodological issues in research on sexual behavior with Latino gay and bisexual men. *American Journal of Community Psychology* 2003;31:281-291.

Zimmerman MA, Ramirez-Valles J, Suarez E, de la Rosa G, Castro MA. An HIV/AIDS prevention project for Mexican homosexual men: an empowerment approach. *Health Educ Behav* 1997;24:177-190.

Appendix A. Correcting the Variance of the Intervention

Effect in Group-Randomized Trials when Only the Variance

Appropriate to an Individually-Randomized Trial Is Available

Correcting the Variance of the Intervention Effect in Group-Randomized Trials when Only the Variance Appropriate to an Individually-Randomized Trial Is Available

DRAFT DATE   July 28, 2009

Word count: 7702 (main text only, excluding appendices, tables, refs, etc.)

Running header: Transformations Between Count and Dichotomous Data

Wayne D. Johnson

e-mail: wjohnson@cdc.gov

Summary

Published analyses for group-randomized trials often do not account for intraclass

correlation (ICC), the propensity for similar responses among members within groups.

Reports of such studies often include only the sample size and either the proportion

experiencing the outcome of interest (for dichotomous variables) or the mean and

standard deviation (for count or continuous variables) for each treatment condition

without regard to groups. Subsequent meta-analyses must rely on summary statistics to

estimate the variance of the intervention effect but must also account for the effects of

group randomization. When only one group has been randomized to each treatment

condition (which occurs frequently in public health), ICC cannot be estimated from the

data. Here we derive and test the factors necessary to adjust variances for ICC due to

group assignment by comparing hypothetical ANOVA tables for individual-randomized

trials and group-randomized trials. For a partly naïve analysis (which correctly accounts

for pre-to-post correlation within member $\hat{r}_{yy(m)}$) of the nested cohort design, and for

posttest-only and nested cross-sectional designs, Donner's variance inflation factor

$V\hat{I}F = 1 + (m-1)I\hat{C}C$ yields a satisfactory approximation of the correction factor if ICC is

small. For a completely naïve analysis of the nested cohort design that does not account

for pre-to-post correlation within member, a satisfactory approximation of the correction

factor is $V\hat{I}F \times \left(1 - \hat{r}_{yy(m)}\right)$. We demonstrate the impact of this correction in empirical

studies and meta-analysis of HIV prevention research. These correction factors are useful

for meta-analysis, power calculations, and studies where only one or a few groups have

been randomized to each treatment condition.

1. BACKGROUND

Interventions directed to communities or other socially intact groups are essential to health promotion. The nature of some interventions requires that they be delivered in the context of social groups. Interventions delivered to communities may also include scaled-up effectiveness trials of individual-level interventions. The variance of the intervention effect in such studies is influenced by intraclass correlation (ICC), the propensity for members of a group to give similar responses.

Because such studies often have low statistical power, it is particularly important to summarize and compare information from multiple trials in meta-analysis. At least two groups per condition are required for the estimation of ICC within a study, but many if not most studies do not meet even this minimal criterion. Even when two or more groups are assigned to each treatment condition so that a valid analysis could be performed, research teams may not be aware of the need to analyze the data by a method (e.g., general or generalized linear mixed models) that accounts for groups as units of assignment. A procedure for correcting the variance estimate is needed so that studies which have not been (or could not be) correctly analyzed in primary reports can still be included in subsequent meta-analysis. Valid estimates of variance accounting for ICC are required for this process because the reciprocal of the variance is used to weight the studies.

It is known that the variance of a mean or proportion sampled in groups can be multiplied by a variance inflation factor $VIF = 1 + (m-1)ICC$ to account for ICC [Donner 1981, Donner 1984]. While different factors have been proposed to adjust for group randomization [Rooney & Murray, Hsieh], to our knowledge their application to

the estimating the variance of intervention effects has not been explicitly examined in the context of meta-analysis with theoretical derivations, testing by simulation, and evaluation by comparison with the correctly estimated variance in empirical examples.

Correction for intraclass correlation influences a meta-analysis because the contribution of each study to the summary effect is weighted by the reciprocal of the variance. Thus the meta-analytical weight of each study is reduced by the same factor by which the variance is increased. The correct estimate of a study's weight is the reciprocal of the variance from the correct mixed model that accounts for group membership. In contrast, the naïve weight is the reciprocal of the variance obtained from an analysis that ignores ICC, and thus is subject to an increased type 1 error rate if ICC is greater than zero.

The purpose of this manuscript is to derive and evaluate expressions for correction factors for the variance of the intervention effect when complete data are not available, but when certain variance components or correlations can be estimated or inferred from other studies. We then illustrate application of the correction factors using published studies.

The manuscript is organized as follows: First, we derive correction factors under the linear or additive model by comparing the variance components for a hypothetical individual-randomized trial with those for a hypothetical group-randomized trial, and we verify with simulated data that the product of the resulting correction factors times the naïve variance of the intervention effect does indeed yield the variance estimated from the proper analysis that accounts for the group randomization (section 2). Second, to examine empirically and compare the performance of these correction factors with the

traditional VIF for correcting the variance of the intervention effect on a dichotomous outcome in logistic models, we conduct Monte Carlo simulations in which we generate (section 3) and analyze (section 4) simulated data under a series of multiplicative models. Finally we illustrate the application to empirical studies (section 5) and meta-analysis (section 6) of community-level intervention studies for HIV prevention among men who have sex with men, and discuss these findings and applications (section 7).

2. DERIVATION FOR LINEAR MODELS

In the following section and the appendix we derive the correction factors under linear models for three designs – the posttest-only design, the nested cross-sectional design, and the nested cohort design – by comparing variance components from mixed models (which are appropriate for group-randomized trials) with those obtained from the naïve fixed effects models (which are appropriate only for individual-level randomized trials). We follow the notation used by Murray [1998]. Consider a balanced design in which g groups (communities, hospitals, schools, classrooms, social networks, etc.) are allocated to each of c treatment conditions (we will consider only two: experimental and control). Each group includes m members (students, patients, residents, etc.). We will consider examples only with equal numbers of members per group and equal numbers of groups per treatment condition. We label the number of members per treatment as $n = gm$. The entire study has $2g$ groups and $2n = 2gm$ individuals.

We use the subscript N when necessary to distinguish components specific to the naïve model (a fixed effects analysis ignoring group membership) and C for the correct model (a mixed model including group membership as a random effect). A total of g communities are assigned to receive the intervention and g to a wait-list control

condition. In each community, *m* sexually active participants respond to a survey after the intervention is delivered in the experimental treatment communities. Relevant examples might include published articles evaluating an HIV prevention program in different communities. Published results frequently include only the mean and standard deviation of the proportion of intercourse occasions when condoms were used and the number of subjects in each treatment condition, not for each randomized group within each condition. If the individual-level data or at least the group-level means and standard deviations are not available, an external estimate of ICC must be substituted.

The Individual- and Group-Randomized Posttest-Only Designs

In the simplest *individually* randomized control trial design, the Posttest-Only Control Condition design [Murray pg 355], the naïve model incorporates an independence assumption, as though each person had been individually assigned to either the experimental or control condition. This naive model is implicitly:

$$Y_{i:l} = \mu + C_l + \boldsymbol{\varepsilon}_{i:l}$$

Here, $Y_{i:l}$ represents the outcome for the *i*-th person assigned to the *l*-th treatment condition where $i = 1, 2, \ldots, m$ and $l = 0, 1$ (0 for the control condition and 1 for the experimental intervention), μ is a fixed effect representing the overall mean when other fixed effect values are equal to zero. C_l is a fixed effect value that represents the effect of the intervention. In this *fixed effects* model, $\boldsymbol{\varepsilon}_{i:l}$ is the random error for the *i*-th person in the *l*-th treatment condition (indicated by bold text). For this simple ANOVA model, we make three assumptions: the errors are independent, identically distributed, and homoscedastic, that is, the variances are equal at all levels of the explanatory variables. In

the applications and to calculate certain statistics, one may further assume normality of the errors.

The quantity of primary interest is the *intervention effect* Δ, which is estimated by the difference between the mean, $\overline{Y}_1$, among those assigned to condition 1 (intervention) and the mean, $\overline{Y}_0$, among those assigned to condition 0 (comparison):

$\Delta = C_1$ which is estimated by the difference in group means $\hat{\Delta} = \overline{Y}_1 - \overline{Y}_0$

In contrast, in the appropriate *mixed* model for the simple *group*-randomized trial design [Murray pg 361; Neter et al pg 981-2] a random effect for each *group* (e.g., school, hospital, neighborhood, etc) is included to account for group membership. This model will be referred to as the "correct" model. It can be expressed as:

$Y_{i:k:l} = \mu + C_l + G_{k:l} + \varepsilon_{i:k:l}$

Here, $Y_{i:k:l}$ represents the outcome for the i-th person in the k-th group that is assigned to the l-th treatment condition; μ represents the overall mean when other fixed effect values are set to zero;, C_l is a fixed effect of treatment ; $G_{k:l}$ is a random effect distributed as $G_{k:l} \approx N\left(\sigma_{g:c}^2\right)$ that represents the effect of membership in the k-th group. The random effect $\varepsilon_{i:k:l}$ represents the deviation of the i-th individual's response from the k-th group mean, nested within the l-th treatment condition mean.

For this mixed effects ANOVA model we assume the observations are identically distributed and homoscedastic. But the random group effect allows for a correlation structure reflected by the intraclass correlation, which for this design is

$$ICC_{m:g:c} = \frac{\sigma_{g:c}^2}{\sigma_{g:c}^2 + \sigma_e^2} \ .$$

The quantity of primary interest is again the intervention effect which is estimated by the difference in group means:

$$\hat{\Delta} = \overline{Y}_1 - \overline{Y}_0$$

The expected sums of squares based on the correct model (indicated by the subscript C) [Murray, pg 133] are:

$$E[SSE_C] = 2g(m-1)\sigma_e^2$$

$$E[SS(G:C)] = 2(g-1)\left(\sigma_e^2 + m\sigma_{g:c}^2\right)$$

However the naïve model does not account for group membership, so the sums of squares associated with group membership are erroneously left to fall into the error term of the naïve (N) model. Thus the expected value of the error sum of squares under the naïve model is the sum of the expected value of the error sum of squares under the correct model plus the expected value of the sum of squares for group membership:

$$E[SSE_N] = E[SSE_C] + E[SS(G:C)]$$

Under the naïve (mis-specified) model the variance of the intervention effect, in expectation, is (erroneously) taken to be:

$$\sigma_{\Delta,N}^2 = \frac{E[SSE_N]}{n(n-1)} = \frac{2\left[(n-1)\sigma_e^2 + m(g-1)\sigma_{g:c}^2\right]}{n(n-1)}$$

However, the correct variance of the intervention effect [Murray pages 133-134] is:

$$\sigma_{\Delta,C}^2 = \frac{2E[MS_{g:c}]}{n} = \frac{2\left(m\sigma_{g:c}^2 + \sigma_e^2\right)}{n}.$$

Now we seek a correction factor (cf) such that $cf_{posttest-only} \times \sigma_{\Delta,N}^2 = \sigma_{\Delta,C}^2$. Thus,

$$cf_{posttest-only} = \frac{\sigma_{\Delta,C}^2}{\sigma_{\Delta,N}^2} = \frac{(n-1)\left(m\sigma_{g:c}^2 + \sigma_e^2\right)}{(n-1)\sigma_e^2 + m(g-1)\sigma_{g:c}^2} \qquad 1$$

Substituting the expression for the ICC (above) into equation 1 gives:

$$cf_{posttest-only} = (n-1)\frac{\dfrac{\sigma_e}{\sigma_{g:c}^2+\sigma_e^2}+\dfrac{m\sigma_{g:c}^2}{\sigma_e^2+\sigma_{g:c}^2}}{\dfrac{(n-1)\sigma_e^2}{\sigma_{g:c}^2+\sigma_e^2}+\dfrac{m(g-1)\sigma_{g:c}^2}{\sigma_{g:c}^2+\sigma_e^2}} = \frac{(n-1)\left[1+(m-1)ICC_{m:g:c}\right]}{(n-1)\left(1-ICC_{m:g:c}\right)+m(g-1)ICC_{m:g:c}}$$

Now because $VIF_{m:g:c} = 1+(m-1)ICC_{m:g:c}$ [Murray, page 231], the correction factor can be expressed as:

$$cf_{posttest-only} = \frac{(n-1)VIF_{m:g:c}}{n-VIF_{m:g:c}}$$

We estimate $\sigma_{g:c}^2, \sigma_e^2, ICC_{m:g:c}$, and $VIF_{m:g:c}$ by substitution of the appropriate estimates and sums of squares (summarized in Table 1).

Thus the correct variance of the intervention effect that accounts for correlation of responses within groups can be estimated as the product of a factor involving n and the estimated VIF times the naïve variance estimate. If $ICC_{m:g:c}$ is small, then the correction factor approximately equals the VIF.

Derivations of correction factors for the other group randomized designs are given in the Appendix, and shown in table 2. If the outcome of interest is continuous as in the derivations above, the general linear mixed model is used to measure the intervention effect as the difference between means of a normally distributed outcome.

The general linear mixed model has also been found to yield satisfactory estimates of the intervention effect and its variance when the outcome of interest is *dichotomous* and the intervention effect is to be measured as a difference between proportions (Hannan and Murray 1996). We further examined the effects of these correction factors on simulated data in 100 iterations for each of the three designs,

generating dichotomous outcomes under an additive model (data not shown). The

additive models for generating these data are shown in Table 3. In each case we

considered, the variance estimated by the properly specified, generalized linear mixed

model (GLIMMIX) was accurately obtained within rounding error by multiplying the

naïve variance by the correction factor.

Thus these correction factors, which were derived under the assumptions of

identically distributed error terms, appropriate for many mixed linear models with a

normally distributed (typically continuous) outcome, were also fully successful for the

situations considered in correcting the naïve variance of the intervention effect in mixed

linear models with a dichotomous outcome, where the intervention effect is measured as

a difference in proportions. However it remains to be shown whether the same correction

factors work appropriately for logistic models, where the outcome is dichotomous and the

intervention effect is modeled not as a difference but as an odds ratio.

3. GENERATING SIMULATED DATA UNDER MULTIPLICATIVE MODELS

MODELS FOR LOGISTIC ANALYSIS

The correction factors described above were derived by comparison of expected

sums of squares under assumptions for a linear ANOVA model. But perhaps the more

frequently used method for estimating intervention effects and variances for dichotomous

outcomes is logistic analysis. In this case the metric of effect is not the difference

between proportions but the odds ratio. Critical to our situation, sums of squares are not

used in the logistic model. Therefore it remains to be shown whether the derived

correction factors continue to work well when the data are generated under a

multiplicative model and either a naïve or a correctly specified mixed logistic regression

model is used for the analysis. It also remains to be shown how ICC can be estimated for mixed logistic models.

To evaluate this issue empirically we conducted simulations, generating and analyzing binary data using high, medium, and low estimates for background prevalence β_0, secular trend β_2, and intervention effect β_3, as well as for random effects as appropriate to each study design. Multiplicative models for the three designs are shown in Table 4. For each design, the intercept β_0 estimates the logit of the background prevalence of the outcome, and β_3 estimates the log of the odds ratio representing the effect of the intervention. The background prevalence can be estimated by the inverse logit function $(\exp \beta_0)/[1+(\exp \beta_0)]$, and the intervention effect by $\exp \beta_3$ where exp is the exponential function. In the nested cross-sectional and nested cohort designs, the difference between intervention and comparison groups at baseline is estimated by $\exp \beta_1$ and the secular trend (the change across time in the comparison group) by $\exp \beta_2$.

We used data from previous meta-analyses and empirical studies of HIV prevention for men who have sex with men (MSM) to select a plausible range of parameters for the simulations (Table 5). We identified 11 HIV prevention trials where social groups or communities were the unit of assignment to treatment status. Six of these studies used a nested cross-sectional design and five used a nested cohort design. Of the six cross-sectional studies, three assigned at least two communities to each of the two treatment conditions. ICC estimated from the data in these studies ranged from zero to .013. Pre-to-post correlation for groups ranged from .4 to .7 with an average of .6.

Of the five cohort studies, only one assigned more than one group to each of the two treatment conditions. The ICC estimated from this study was .05, and pre-to-post

correlation for groups was .02. This study was somewhat different from the others in that it was conducted in Russia and Bulgaria (as opposed to the US or UK), the groups were small social networks with an average size of 4 (as opposed to gay gyms or larger gay communities), and 52 units were randomized (as opposed to only 2 or 3 in the other cohort studies). Pre-to-post correlation for members ranged from .4 to .7 with an average of .5.

To generate simulated data, we assumed three groups in each of the two conditions (intervention and comparison) for all designs. We assumed 300 members per group for the posttest-only and nested cross-sectional designs, and 100 for the nested cohort design (because larger sample sizes required excessive computation time). For each design, we first conducted one set of simulations using medium values for all parameters (the base case). We then conducted analogous simulations under different scenarios substituting the alternative high and low values for each parameter. For the fixed effects we selected high, medium, and low values of 0, -1, and -2 for β_0; +.35, 0, and -.35 for β_2; and 0, -.35, and -.7 for β_3. We assumed the average baseline difference β_1 was zero. We then selected values for the random effects that yielded correlations similar to those observed in the empirical studies. For all three designs we selected high, medium, and low values of .5, .17, and .06 for γ_1. For the nested cross-sectional design we selected .7, .2, and .1 for γ_2. For the nested cohort design we selected .2, .1, and 0 for γ_2; and 3, 2, and 1 for γ_3.

We performed separate series of simulations for three designs. For the cohort design we considered two naïve analyses. The partly naïve analysis accounts for member as a random effect and thus accounts for $\hat{r}_{yy(m)}$, but still ignores ICC and $\hat{r}_{yy(g)}$. The

completely naïve analysis for the cohort design ignores all 3 correlations (ICC, pre-to-post for groups $\hat{r}_{yy(g)}$ and pre-to-post within member $\hat{r}_{yy(m)}$).

Corrected variances were obtained by multiplying the naïve variance by the traditional VIF and by the new CF; corrected weights were the reciprocals of these variances. For the completely naïve analysis under the cohort design, the variance was further multiplied by a factor of $1 - \hat{r}_{yy(m)}$ where $\hat{r}_{yy(m)}$ is the observed pre-to-post correlation within member.

Finally we assessed the precision of the naïve estimate and each of the corrected estimates by calculating the mean squared error (MSE) between each approximation and the true weight. MSEs closer to zero indicate better estimates. We took the average of all informative iterations for each scenario, excluding iterations where any necessary model did not converge or where ICC=0, since results are non-informative in such cases. We did not allow negative variance components, which would have required excessive computation time. As a result, variance components are biased in a positive direction and will not precisely correspond to our input parameters, but this has no bearing on our investigation of the effectiveness of correction factors or VIF in adjustment of the variance in the remaining iterations. For each design, we present values for fixed and random effects, correlations, VIF, CF, and true, naïve, and corrected estimates of weight and MSE averaged across all convergent iterations where ICC>0.

4. SIMULATIONS USING LOGISTIC MODELS

To analyze the simulated data generated under the multiplicative models described above, we applied the Generalized Linear Mixed Model (Wolfinger and O'Connell 1993, Breslow and Clayton 1993). In this approach, "pseudo-data" are

iteratively constructed by Taylor series expansion to permit modeling of the dichotomous outcome. Covariances and the residual error are then estimated by the pseudolikelihood method in a manner analogous to that used for linear data as described in Table 1.

If the estimation is based on residual likelihood, and the expansion locus is the marginal mean of the random effects (as specified by METHOD=RMPL in SAS), then the point estimate of the intervention effect in a balanced design is the same as the point estimate obtained from the naïve logistic model. Thus, estimates of the log odds ratio from the collapsed contingency table with no information concerning group membership are identical to estimates from the SAS GLIMMIX procedure using the RMPL (residual marginal expansion pseudo-likelihood) technique, and we need be concerned with correcting only the variance.

Variance components from Proc Glimmix in SAS are presented on the scale of the link function used for the generalized linear mixed model (Murray page 239). This means that values presented for the residual and for member must be multiplied by the variance function, which in the case of the binomial distribution is $\bar{p}(1-\bar{p})$ where $\bar{p}$ is the overall prevalence of the outcome, while values presented for group nested within condition and for time by group nested within condition must be multiplied by the square of the variance function (Table 6). These values can then be used to estimate ICC, $r_{yy(g)}$, and $r_{yy(m)}$ by the formulas in Table 2. Correlations estimated by these methods are similar to those estimated from linear models. This raises the question of how well the estimates of correlation from the linear model may perform for correcting naïve estimates of the variance of the log odds ratio. Thus we will consider both the correlations obtained from

the linear model and the correlations obtained from the logistic model for estimating the

VIF and CF, when we compare MSEs for various methods of estimation.

A further set of options is raised concerning the estimation of $\bar{p}$. This value can

be estimated in at least three ways: the overall average proportion, the inverse logit of the

average of the four logits of the proportion in each treatment condition at each time, or

the inverse logit of the average of the 12 logits of the proportion in each group within

each treatment condition at each time. We use these 3 estimates of $\bar{p}$ to generate 3

estimates of ICC based on logistic model parameter estimates. We label these estimates

of ICC in Table 7 as log1, log4, and logX respectively. Finally we use the linear estimate

of ICC and these 3 logistic estimates of ICC to generate 4 estimates of VIF and 4

estimates of CF.

Thus for each design, we want to compare the performance of the newly derived

CF versus the traditional VIF, using estimates of correlations based on linear versus

logistic models, and for logistic models, with estimates of the scaling factor based on the

overall prevalence, four prevalences (2 conditions at 2 times), or 12 prevalences (3

groups within each of 2 conditions at 2 times).Performance will be compared by

manipulating each of the 3 characteristics (CF vs VIF, linear vs logistic correlations, and

type of prevalence estimate for the scaling factor in logistic correlations) in an effort to

obtain minimal values of MSE.

For each of the study designs we present the number of successful (convergent)

iterations, baseline prevalence (for the intervention and comparison groups combined,

since we assume that they do not differ at baseline), the intervention effect, the ICC, and

the resulting naïve and correct meta-analytic study weights (Table 7). Substantial

numbers of iterations were excluded due to non-converging logistic models (where no results were available) or ICCs of zero (where the results under the naïve model were identical to the results of the correct model, and thus provided no useful information for our question).

Post-test only and nested cross-sectional designs. The average value of about -1 for β_0 indicates an average comparison prevalence of $e^{-1}/(1+e^{-1}) = 27\%$. The average value of about -.34 for β_3 indicates an average odds ratio of $e^{-.34} = .71$ for the intervention effect. The average ICC was about .01 and the average VIF and CF were about 4 by all four methods (linear and the three logistic methods).The average naïve weight was more than twice the average true weight. The average weights corrected by VIF approached the true estimate of weight as the estimate of ICC progressed from the linear value through the more precise logistic estimates. A similar pattern was observed among weights corrected by CF. Estimates were slightly closer to the correct value when obtained by way of the traditional VIF for the post-test only design, and by way of the newly derived CF for the nested cross-sectional design.

MSE for the naïve estimates of weight were 40 to 80 times the average weight. MSEs confirm the impression that corrected values obtained by way of VIF were slightly more accurate than those by way of CF for the posttest-only design, but the reverse for the nested cross-sectional design. Estimates based on group-specific $\bar{p}$ were very slightly more accurate than those by way of the overall $\bar{p}$.

Nested cohort design. The average β_0 of about -0.6 indicates an average baseline prevalence (in the intervention and comparison conditions combined, because we assume the baseline difference between the two conditions is zero) of 35%. The average β_3 of -.22

indicates an odds ratio of .80 for the intervention effect. The partly naïve analysis of the nested cohort design includes somewhat fewer iterations than the completely naïve analysis because the required mixed logistic model (accounting only for $\hat{r}_{yy(m)}$ and not ICC or $\hat{r}_{yy(g)}$) did not converge in several iterations.

Fixed parameter estimates and the correlations changed little from the average values observed under the post-test only and nested cross-sectional designs. The average ICC was about .01, yielding average VIF and CF of about 2.2 by all 4 methods for the partly naïve analysis and 1.4 (including a factor of $1 - \hat{r}_{yy(m)}$ for VIF) for the completely naïve analysis.

For the partly naïve analysis, the average naïve weight of about 27 was almost twice the average true weight of about 15; the MSE of 180 was about 12 times the average true weight. For the partly naïve analysis, correction by way of the traditional VIF yielded somewhat more precise results than by the newly derived CF. MSEs were reduced by a factor of about 90 (180/2) for corrections based on the traditional VIF and linear correlations. MSEs for corrections based on the traditional VIF and logistic correlations were even more precise by another factor of 100 (2/.02). There was a very slight advantage of estimating $\bar{p}$ from the group-specific $\bar{p}$ over the overall $\bar{p}$ as the scaling factor.

For the completely naïve analysis, the average naïve weight of about 16 was not very different from the average true weight of about 15 because the apparent precision incorrectly gained by neglecting ICC was approximately balanced by an incorrect loss of precision due to neglecting $\hat{r}_{yy(m)}$. Nevertheless the MSE of about 46 for the naïve weight under the completely naïve model was about 3 times the average true weight. Correction

by way of the newly derived CF yielded somewhat more precise results than by the traditional VIF. MSEs were reduced by a factor of about 20 (46/2.2) for corrections based on the newly derived CF and linear correlations; MSEs for corrections based on the newly derived CF and logistic correlations were even more precise by another factor of over 100 (2.2/.02). Again there was a slight advantage of estimating $\bar{p}$ from the group-specific $\bar{p}$ over the overall $\bar{p}$ as the scaling factor.

Summary. Thus in simulations with logistic models, CF performed slightly better than $\hat{VIF}$ for the nested cross-sectional design and the completely naïve analysis of the nested cohort design. In contrast, $\hat{VIF}$ performed slightly better for the partly naïve analysis of the nested cohort design. Performance of CF and $\hat{VIF}$ were almost indistinguishable for the post-test only design.

Estimates based on correlations from the logistic model using $\bar{p}$ as the scaling factor were substantially more accurate than estimates based on correlations from the linear model by a factor of 15 to 150. Estimates based on correlations from the logistic model using the inverse logit of the average of the logits of the 12 time-by-group proportions as the scaling factor were usually more accurate than those using $\bar{p}$ by a factor of .8 (favoring $\bar{p}$) to 3.0 (favoring logX). But all correction methods provided a substantial improvement over the naïve estimate, by factors of 72 to over 30,000.

Thus among the three characteristics involved in the correction process, (CF vs VIF, linear vs logistic correlations, and type of prevalence estimate for the scaling factor in logistic correlations), the use of correlations estimated from logistic models rather than from linear models was the most critical to performance as measured by minimal MSEs. For the post-test only design for example, MSEs were about 0.1 to 0.2 when correlations

estimated from logistic models were used, compared to about 3.5 when correlations estimated from linear models were used.

The choice concerning which estimate of $\bar{p}$ to use in the scaling factor led to only very small differences in MSE. Even so, those small differences were mostly in the expected direction. Estimates of $\bar{p}$ as the inverse logit of the average of the 12 logits of the proportion in each group within each treatment condition at each time yielded the smallest MSE, followed by estimates of $\bar{p}$ as the inverse logit of the average of the four logits of the proportion in each treatment condition at each time, followed by estimates of $\bar{p}$ as the overall average proportion.

In practical applications, both the choice between correlations derived from linear and logistic models, and the choice among methods for estimating the scaling factor, are moot. All three correlations ($I\hat{C}C$ and, if applicable, $\hat{r}_{yy(g)}$ and $\hat{r}_{yy(m)}$) must be estimated very roughly based on results of other studies. Given the uncertainty of the three correlations, the difference in performance between $V\hat{I}F$ and CF is likely to be negligible in most empirical applications.

5. APPLICATION TO EMPIRICAL STUDIES

Next we examine the effect of applying correction for ICC to the weight of 11 available studies of behavioral HIV prevention for MSM (Table 8). For all but the far right column of this table we collapse multiple data collection waves into a single baseline and a single followup. Given the small difference observed above between VIF and the correction factors, we apply the simpler factor $V\hat{I}F \times \left(1 - \hat{r}_{yy(m)}\right)$ to completely naïve analyses of the nested cohort design and $V\hat{I}F$ alone in all other cases.

For each study, we estimate the VIF as $1 + (m-1)$ICC assuming that ICC = .005 or ICC = .03. Then we show the estimates of weight, first assuming ICC = 0, and then after applying the estimates of VIF assuming that ICC = .005 or ICC = .03. Next we show the estimates of weight applying the VIF for those studies where ICC could be estimated from the data. Finally in the rightmost column, we show the estimate of weight for a complete analysis of the individual data where multiple waves are not collapsed, and time is treated as a continuous variable in studies with more than two waves of data collection (measuring the proportion of total study period elapsed) rather than dichotomizing time to represent either before or after intervention.

Samples and cohorts included 9118 responses at baseline and 8880 at followup. Almost two-thirds ($k = 7$) of the studies included only one group or cluster in one or both conditions, so the usual mixed model analysis that estimates ICC from the raw data would not be possible. Table 8 shows that the impact of correction for ICC increases not only with the assumed value of ICC but also with the average group size. Since at any given ICC, VIF and CF depend on the average cluster size, those studies with the largest cluster size are penalized the most by variance inflation. For example, at an assumed ICC of .005, the naïve variance for the two largest studies (with an average of about a thousand members per group) must be multiplied by a factor of 6 or 7. Four studies with an average of 200 to 300 members per group had correction factors of 2 to 3. The correction factor was about 1.5 for studies with around 100 members per group, and little correction was necessary for studies with fewer than 50 members per group.

Similarly, at a given cluster size m, the corrected variance increases substantially, and thus the weight decreases, with increasing assumed values for ICC. For example, if

ICC is assumed to be zero, the study with the largest N of 2324 has a weight of 59.7. But even at a moderate ICC of .005 this weight is reduced to 8.8 and at a high ICC of .03 the weight is 1.7, barely larger than the study with a total sample size of only 54. By contrast, the weight of the study with the lowest average group size m changes little from 9.6 if ICC = 0 to 8.8 if ICC = .03.

Estimates of ICC could be obtained from the data in four studies. Results of this adjustment ranged from no impact on Elford 2002 (where ICC was estimated from the data to be zero) to the greatest impact on ACDP 1999 (where ICC was estimated from the data to be .013). Even though the highest ICC estimate of .05 was observed for Amirkhanian 2005, the average group size m was smallest in this study and the resulting VIF was small.

Complete analyses based on published results or raw data were available for the same 4 studies. These analyses accounted not only for the internal estimate of ICC but also for design features unique to each study such as multiple waves of followup data collection, matched pairs of intervention and comparison groups, and unequal number of members per group and across time. This rightmost column illustrates the additional advantage of multiple waves of data collection, independent of the intraclass correlation, with the weight being more than doubled in one study.

6. META-ANALYSES OF EMPIRICAL STUDIES

We then conducted and compared meta-analyses of the 11 studies under each of the assumptions described in Section 5. The total weight in fixed effects meta-analyses of these studies decreases dramatically with increasing assumed values of ICC, from a weight of 225 if ICC is assumed to be zero to only 31 if ICC is assumed to be .03 (Table

9). The impact of this overall decrease in weight is an increase in the variance and thus the width of the fixed effects confidence interval from (.66, .85) at ICC=0 to (.41, .83) at ICC=.03.

The change in individual study weights relative to each other can affect not only the width of the confidence interval but also the point estimate itself. In our case, the sharp decrease in weight for the study with the least favorable overall effect (Elford 2002) combined with a relatively stable weight for the study with the most favorable effect (Amirkhanian 2005) results in progressively *more* favorable point estimates from an odds ratio of .75 when ICC is assumed to be zero to an odds ratio of .58 when ICC is assumed to be .003. Application of study-specific ICC estimates and complete analyses returns the point estimates toward the more moderate range of .72 to .73, and the confidence intervals to a moderate width (.59 to .87 for complete analyses).

The Q statistic is distributed as chi-square and tests for heterogeneity among studies. The small values of the Q statistic in the first 5 rows indicate that given the assumptions within each meta-analysis, the effect sizes are relatively homogeneous. The additional component of variance tau which reflects that homogeneity is therefore small for each meta-analysis, and the net results of random effects meta-analyses are similar or identical to the fixed effects meta-analyses for the first 5 sets of assumptions.

Finally, if the correction factors were not available, we would have been obliged to exclude the 7 studies for which an internal estimate of ICC was not available. In this case, heterogeneity is quite significant ($Q_3 = 9.4$, p=.02) necessitating use of the random effects model. The point estimate (OR = .67) is a little stronger than most other summary estimates in the table, but the confidence interval includes the null value of 1.0, indicating

a non-significant summary effect. Even worse, almost two-thirds of the studies would be lost, leaving little opportunity to investigate the cause of the heterogeneity.

7. DISCUSSION

We have derived correction factors to account for assignment of existing social or other clusters to treatment status in group randomized controlled trials. The correction factors were derived for linear models, where the outcome is generally assumed to be continuous with identically distributed errors, and the intervention effect is measured as a mean difference between conditions. We showed by simulations that the correction factors also perform well in logistic models where the intervention effect is measured as the odds ratio.

When ICC is small, the correction factors are approximately equal to the VIF, except for the case of the completely naïve model for the cohort design, in which case the correction factor is approximately $V\hat{I}F \times \left(1 - \hat{r}_{yy(m)}\right)$. This simplified approach using only the VIF, or the $V\hat{I}F \times \left(1 - \hat{r}_{yy(m)}\right)$, instead of the full correction factors also performed well in simulations.

Appropriate corrections to the meta-analytical weights are necessary for studies where existing groups are the unit of assignment. In our empirical examples, some study weights were reduced by factors of about 6 or 7 even at a modest ICC of .005. The fixed effects summary estimate actually moved farther from the null as assumed values of ICC increased. This apparent paradox occurred because the VIF happened to be greatest for two studies that had the least favorable effects.

This change in weight can affect meta-analysis in at least four ways. First, the overall reduction in the weight of available evidence could lead to an increase in the

variance to the extent that summary results that would be statistically significant in a naïve analysis are no longer significant. Second, the greater reduction in weight from studies with a large number of members, while other studies with a large number of groups are less affected, could change the magnitude of the summary intervention effect from a value reflecting mainly the former to a value reflecting mainly the latter. Third, accounting for ICC makes weights smaller and tends to make weights more similar across studies. Since small outliers introduce less heterogeneity than large outliers, the difference (if any) between results of fixed and random effects analyses may decrease when ICC is accounted for. Finally, within a meta-analysis that also includes studies where individuals are the unit of assignment, the reduction in weight for studies where groups are the unit of assignment yields a summary result that more accurately reflects the proportion of information that comes from these studies. Without correction for ICC, the intervention effects of studies where groups are the unit of assignment would be overrepresented in the summary result.

Outside of meta-analysis, this correction process is also useful for individual studies where only one cluster has been assigned to each treatment condition. Confidence intervals and statistical significance can be presented for a range of assumed values of ICC from similar studies, for example zero (no correction), .005, and .030. Even when two or three groups have been assigned to each condition so that ICC can be estimated, the estimate has a wide confidence interval and it may be prudent to use these formulas to consider a range of possible values for ICC.

Setting aside the issue of ICC for a moment, it is also remarkable that the correct estimates of weight of about 15 under the cohort design were not much less than the

correct estimate of about 18 under the cross-sectional design, even though the sample size (N=1800) assumed for the cross-sectional design at each time point was three times as great as for the cohort design (N=600). This result illustrates in part the futility of large cluster sizes in group-randomized trials. If we assume that ICC is .005 in both cases then the VIF, which is $1+(m-1)ICC$, increases from about 1.5 when $m = 100$ to 2.5 when $m = 300$. The cohort design further benefits in statistical power from the factor of $\left(1-\hat{r}_{yy(m)}\right)$ which reflects correlation within member across time. If that correlation is 0.5, then the weight under a cohort design is approximately doubled (and the variance decreased by half) compared to a cross-sectional design with the same sample size and ICC.

Our derivation and simulations do not address matched or stratified designs, unequal cluster sizes, multiple waves of followup data collection, analyses of followup data with adjustment for baseline conditions as a covariate, or individually-randomized group treatment trials (Diehr 1995, Feng 1999, Kerry 2001, Pals 2008). Further, this correction process adjusts only the variance and associated statistics such as standard error and meta-analytical weight. It cannot address concerns regarding the validity of the point estimate when only one cluster has been randomized to each treatment condition. Such point estimates are subject to bias because an insufficient number of units have been randomized to account for different trajectories that might have occurred even in the absence of the intervention. This bias can be reduced by combining multiple studies within meta-analysis.

Although group-randomized trials often have little power to detect a statistically significant intervention effect when ICC is properly accounted for, meta-analysis of several such studies can provide more stable and valid estimates. If an external estimate

or likely range of ICC is available, correction factors for linear and logistic models are straightforward and easy to apply. The VIF (or $\hat{VIF} \times \left(1 - \hat{r}_{yy(m)}\right)$ in the case of the completely naïve model for the cohort design) performs well in the range of correlations we have examined. This correction can also be used to estimate variance even in the case of only one or a small number of clusters per treatment condition, which is common in community-level intervention studies, and where estimates of ICC are unstable or cannot be estimated.

Acknowledgments

We are grateful to the authors of the primary studies who provided supplementary data: Yuri Amirkhanian, Graham Bolding, Jonathan Elford, Paul Flowers, Graham Hart, Colleen Hoff, Susan Kegeles, Jeff Kelly, Timothy McAuliffe, Robin Miller, Jonathan Shepherd, and Lisa Williamson. We also thank Sherri Pals for many helpful insights and for reviewing an earlier draft of this manuscript.

References

Breslow NE, Clayton DG. Approximate Inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association*, 1993;88:9–25.

Campbell MK, Grimshaw JM, Steen IN. Sample size calculations for cluster randomized trials. *Journal of Health Services Research and Policy.* 2000;5:12-16.

Cornfield J. Randomization by group: A formal analysis. *American Journal of Epidemiology.* 1978;108:100-102.

Diehr P, Martin DC, Koepsell T, Cheadle A. Breaking the matches in a paired t-test for community interventions when the number of pairs is small. *Statistics in Medicine.* 1995;14:1491-1504.

Donald A, Donner A. Adjustments to the Mantel–Haenszel chi-square statistic and odds ratio variance estimator when the data are clustered. *Statistics in Medicine.* 1987; 6:491-499.

Donner A, Birkett N, Buck C. Randomization by cluster: sample size requirements and analysis. *American Journal of Epidemiology.* 1981;114:906-14.

Donner A. Approaches to sample size estimation in the design of clinical trials – a review. *Statistics in Medicine.* 1984;3:199-214.

Donner A. Some aspects of the design and analysis of cluster randomization trials. *Applied Statistics.* 1998;47:95-113.

Donner A, Klar N. Statistical considerations in the design and analysis of community intervention trials. *Journal of Clinical Epidemiology.* 1996;49:435-9.

Donner A, Klar N. Methods for comparing event rates in intervention studies when the unit of allocation is a cluster. *American Journal of Epidemiology.* 1994; 140: 279-299.

Donner A, Klar N. Issues in the meta-analysis of cluster randomized trials. *Statistics in Medicine.* 2002;21:2971-80.

Donner A, Klar N. Pitfalls of and controversies in cluster randomization trials. *American Journal of Public Health.* 2004;94:416-22.

Feng Z, Diehr P, Peterson A, McLerran D. Selected statistical issues in group randomized trials. *Annual Review of Public Health.* 2001;22:167-87.

Feng Z, Diehr P, Yasui Y, Evans B, Beresford S, Koepsell TD. Explaining community-level variance in group randomized trials. *Statistics in Medicine* 1999;18:539-556.

Hannan PJ, Murray DM. Gauss or Bernoulli? A Monte Carlo comparison of the performance of the linear mixed model and the logistic mixed model analyses in simulated community trials with a dichotomous outcome variable at the individual level. *Eval Rev.* 1996;20:338–352.

Hsieh FY. Sample size formulae for intervention studies with the cluster as unit of randomization. [erratum appears in Stat Med 1997 Jun 15; 16(11):1300] *Statistics in Medicine.* 1988;7:1195-1201.

Hsieh FY, Lavori PW, Cohen HJ, Feussner JR. An Overview of Variance Inflation Factors for Sample-Size Calculation. *Evaluation & the Health Professions.* 2003;26:239-257.

Johnson W, Semaan S, Hedges L, Ramirez G, Mullen PD, Sogolow E. A Protocol for the Analytical Aspects of a Systematic Review of HIV Prevention Research. *J Acquir Immune Defic Syndr.* 2002;30:S62-72.

Johnson W, Hedges L, Ramirez G, Semaan S, Norman L, Sogolow E, Sweat M, Diaz R. HIV Prevention Research for Men Who Have Sex with Men: a Systematic Review and Meta-Analysis. *J Acquir Immune Defic Syndr.* 2002;30:S118-130.

Kerry SM. Unequal cluster sizes in UK general practice: implications for sample size calculations. *Statistics in Medicine.* 2001;20:377-390.

Kish L. *Survey Sampling.* John Wiley and Sons: New York, 1965.

Klar N, Donner A. Current and future challenges in the design and analysis of cluster randomization trials. *Statistics in Medicine.* 2001;20:3729-40.

Klar N, Donner A. The merits of matching in community intervention trials: a cautionary tale. *Statistics in Medicine.* 1997;16:1753-64.

Martin DC, Diehr P, Perrin EB, Koepsell TD. The effect of matching on the power of randomized community intervention studies. *Statistics in Medicine.* 1993;12:329-339.

Moore H, Summerbell C, Vail A, Greenwood DC, Adamson AJ. The design features and practicalities of conducting a pragmatic cluster randomized trial of obesity management in primary care. *Statistics in Medicine.* 2001;20:331-40.

Murray DM. *Design and Analysis of Group-Randomized Trials.* Oxford University Press: New York, 1998.

Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. *Applied Linear Statistical Models.* Irwin: Chicago, 1996.

Pals SL, Murray DM, Alfano CM, Shadish WR, Hannan PJ, Baker WL. Individually randomized group treatment trials: a critical appraisal of frequently used design and analytic approaches. *American Journal of Public Health.* 2008;98:1418-24.

Rosenthal R, Rosnow RL. *Essentials of Behavioral Research: Methods and Data Analysis.* McGraw Hill: Boston, 1991.

Varnell S, Murray DM, Janega JB, Blitstein JL. Design and analysis of group-randomized trials: a review of recent practices. *American Journal of Public Health.* 2004;94:393-399.

Wolfinger R, O'Connell M. Generalized Linear Mixed Models: A Pseudo-Likelihood Approach. *Journal of Statistical Computation and Simulation*, 1993;4:233–243.

Empirical studies of HIV prevention

Amirkhanian YA, Kelly JA, Kabakchieva E *et al*. A randomized social network HIV prevention trial with young men who have sex with men in Russia and Bulgaria. *AIDS.* 2005;19:1897–1905.

CDC AIDS Community Demonstration Projects Research Group. Community–level HIV intervention in 5 cities: Final outcome data from the CDC AIDS Community Demonstration Projects. *American Journal of Public Health.* 1999;89:336–345.

Elford J, Bolding G, Sherr L. Peer education has no significant impact on HIV risk behaviours among gay men in London. *AIDS.* 2001;15:535–538.

Flowers P, Hart GJ, Williamson LM, Frankis JS, Der GJ. Does bar–based, peer–led sexual health promotion have a community–level effect amongst gay men in Scotland? *International Journal of STD & AIDS.* 2002;13:102–108.

Hoff CC, Kegeles SM, Acree M, Stall R, Paul J, Ekstrand M, Coates TJ. Looking for men in all the wrong places . . . : HIV prevention small–group programs do not reach high risk gay men. *AIDS.* 1997;11:829–831.

Kegeles SM, Hays RB, Coates TJ. The Mpowerment Project: A community–level HIV prevention intervention for young gay men. *American Journal of PublicHealth.* 1996; 86:1129–1136.

Kegeles SM, Rebchook GM, Hays RB, Pollack LM. (2002, July). *Staving off increases in young gay/bisexual men's risk behavior in the HAART era* [AbstractMoPeD3549]. Paper presented at the 14th International Conference on AIDS. Barcelona.

Kelly JA, Murphy DA, Sikkema KJ, McAuliffe TL, Roffman RA, Solomon LJ, et al. Randomised, controlled, community–level HIV–prevention intervention for

sexual–risk behaviour among homosexual men in US cities. *The Lancet.* 1997;
350:1500–1505.

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
population: An experimental analysis. *American Journal of Public Health.* 1991;
81:168–171.

Miller RL, Klotz D, Eckholdt HM. HIV prevention with male prostitutes and patrons of
hustler bars: Replication of an HIV preventive intervention. *American Journal of
Community Psychology.* 1998; 26:97–131.

Shepherd J, Weare K, Turner G. Peer–led sexual health promotion with young gay and
bisexual men: Results of The HAPEER Project. *Health Education.* 1997; 97:204–
212.

Table 1. Estimation of covariance parameters necessary for calculating the 3 correlations ICC, $r_{yy(g)}$, and $r_{yy(m)}$.

Component	Estimator	Where
Post only		
σ_g^2	$\hat{\sigma}_g^2$	$SS(G:C) = 2(g-1)(\hat{\sigma}_e^2 + m\hat{\sigma}_{g:c}^2)$
σ_e^2	$\hat{\sigma}_e^2$	$SS(M:G:C) = 2g(m-1)\hat{\sigma}_e^2$
X-sectional		
σ_g^2	$\hat{\sigma}_g^2$	$SS(G:C) = 2(g-1)(\hat{\sigma}_e^2 + 2m\hat{\sigma}_{g:c}^2)$
σ_{tg}^2	$\hat{\sigma}_{tg}^2$	$SS(TG:C) = 2(g-1)(\hat{\sigma}_e^2 + m\hat{\sigma}_{tg:c}^2)$
σ_e^2	$\hat{\sigma}_e^2$	$SS(M:TG:C) = 4g(m-1)\hat{\sigma}_e^2$
Cohort		
σ_g^2	$\hat{\sigma}_g^2$	$SS(G:C) = 2(g-1)(\hat{\sigma}_e^2 + 2\hat{\sigma}_{m:g:c}^2 + 2m\hat{\sigma}_{g:c}^2)$
σ_{tg}^2	$\hat{\sigma}_{tg}^2$	$SS(TG:C) = 2(g-1)(\hat{\sigma}_e^2 + m\hat{\sigma}_{tg:c}^2)$
σ_m^2	$\hat{\sigma}_m^2$	$SS(M:G:C) = 2g(m-1)(\hat{\sigma}_e^2 + 2\hat{\sigma}_{m:g:c}^2)$
σ_e^2	$\hat{\sigma}_e^2$	$SS(MT:G:C) = 2g(m-1)\hat{\sigma}_e^2$

Note: σ_g^2 = group component of variance; σ_{tg}^2 = time by group component of variance; σ_m^2 = member component of variance; σ_e^2 = error component of variance. Adapted from Murray; posttest only pg 133, x-sectional pg 142, cohort pg 181. [temporary note for our reference: it is correct that σ_e^2 for cross-sectional includes a factor of 4, while post-test and cohort have a factor of 2]

Table 2. Correction factors for the variance of the difference between two study conditions in three study designs for group-randomized trials

Study Design	Correction factor	ICC	$r_{yy(g)}$	$r_{yy(m)}$
Posttest-only	$$\dfrac{(n-1)VIF_{m:g:c}}{n-VIF_{m:g:c}}$$	$\dfrac{\sigma_g^2}{\sigma_g^2+\sigma_e^2}$	NA	NA
Nested cross-sectional	$$\dfrac{(n-1)(m-1)VIF_{m:tg:c}\left(1-r_{yy(g)}\right)}{(m-1)\left(n-VIF_{m:tg:c}\right)-(n-1)\left(m-VIF_{m:tg:c}\right)r_{yy(g)}}$$	$\dfrac{\sigma_{tg}^2}{\sigma_{tg}^2+\sigma_e^2}$	$\dfrac{\sigma_g^2}{\sigma_g^2+\sigma_{tg}^2}$	NA
Nested cohort Partly naïve analysis	$$\dfrac{(n-1)(m-1)VIF_{mt:g:c}\left(1-r_{yy(g)}\right)}{(m-1)\left(n-VIF_{mt:g:c}\right)-(n-1)\left(m-VIF_{mt:g:c}\right)r_{yy(g)}-m(g-1)\left(VIF_{mt:g:c}-1\right)r_{yy(m)}}$$	$\dfrac{\sigma_{tg}^2}{\sigma_{tg}^2+\sigma_e^2}$	$\dfrac{\sigma_g^2}{\sigma_g^2+\sigma_{tg}^2}$	$\dfrac{\sigma_m^2}{\sigma_m^2+\sigma_e^2}$
Completely naïve analysis	$$\dfrac{(n-1)(m-1)VIF_{mt:g:c}\left(1-r_{yy(g)}\right)\left(1-r_{yy(m)}\right)}{(m-1)\left(n-VIF_{mt:g:c}\right)-(n-1)\left(m-VIF_{mt:g:c}\right)r_{yy(g)}-m(g-1)\left(VIF_{mt:g:c}-1\right)r_{yy(m)}}$$			

Note: n = average number of members per study condition; m = average number of members per group (nested within study condition); g = average number of groups per study condition; $VIF = 1 + (m-1)ICC$; ICC = intraclass correlation; $r_{yy(g)}$ = over-time correlation at the group level; $r_{yy(m)}$ = over-time correlation at the member level; σ_g^2 = group component of variance; σ_{tg}^2 = time by group component of variance; σ_m^2 = member component of variance; σ_e^2 = error component of variance; NA= not applicable. The partly naïve analysis accounts for member as a random effect and thus accounts for $r_{yy(m)}$, but ignores ICC and $r_{yy(g)}$. The completely naïve analysis for the cohort design ignores all 3 correlations: ICC, $r_{yy(g)}$, and $r_{yy(m)}$.

Table 3. Additive models for the difference in proportions in a group-randomized trial

Design and outcome	Distribution	Model	Random Effects
Post only $Y_{i:k:l}$ given $\boldsymbol{G}_{k:l}$	$\text{bin}(1, p_{i:k:l})$	$p_{i:k:l} = \mu + C_l + \boldsymbol{G}_{k:l}$	$\boldsymbol{G}_{k:l} \sim N(0,\ \sigma^2_{g:c})$
X-sectional $Y_{i:jk:l}$ given $\boldsymbol{G}_{k:l}$ and $\boldsymbol{TG}_{jk:l}$	$\text{bin}(1, p_{i:jk:l})$	$p_{i:jk:l} = \mu + C_l + T_j + TC_{jl} + \boldsymbol{G}_{k:l} + \boldsymbol{TG}_{jk:l}$	$\boldsymbol{G}_{k:l} \sim N(0,\ \sigma^2_{g:c})$ $\boldsymbol{TG}_{jk:l} \sim N(0,\ \sigma^2_{tg:c})$
Cohort $Y_{ij:k:l}$ given $\boldsymbol{G}_{k:l}$, $\boldsymbol{M}_{i:k:l}$, and $\boldsymbol{TG}_{jk:l}$	$\text{bin}(1, p_{ij:k:l})$,	$p_{ij:k:l} = \mu + C_l + T_j + TC_{jl} + \boldsymbol{G}_{k:l} + \boldsymbol{M}_{i:k:l} + \boldsymbol{TG}_{jk:l}$	$\boldsymbol{G}_{k:l} \sim N(0,\ \sigma^2_{g:c})$ $\boldsymbol{M}_{i:k:l} \sim N(0,\ \sigma^2_{m:g:c})$ $\boldsymbol{TG}_{jk:l} \sim N(0,\ \sigma^2_{tg:c})$

Note: σ^2_g = group component of variance; σ^2_{tg} = time by group component of variance; σ^2_m = member component of variance. Adapted from Murray; posttest only pg 133, x-sectional pg 142, cohort pg 181.

Table 4. Multiplicative models for the odds ratio in a group-randomized trial

Design and outcome	Model

Post only

$\text{Logit}[E(Y_{i:k:l} | \gamma_1)]$ $\beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \gamma_{1, k:l}$

X-sectional

$\text{logit}[E(Y_{i:jk:l} | \gamma_1, \gamma_2)]$ $\beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \gamma_{1, k:l} + \gamma_{2, jk:l}$

Cohort

$\text{logit}[E(Y_{ij:k:l} | \gamma_1, \gamma_2, \gamma_3)]$ $\beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \gamma_{1, k:l} + \gamma_{2, jk:l} + \gamma_{3, i:k:l}$

Note:

β_0 = intercept

β_1 = baseline difference between intervention and comparison on the logit scale

β_2 = difference between baseline and followup in comparison condition on the logit scale

β_3 = intervention effect on the logit scale

X_1 = 1 for intervention condition, 0 for control condition

X_2 = 1 for followup, 0 for baseline

X_3 = 1 for intervention condition at followup, 0 for all others

$\gamma_{1, k:l} \sim N(0, \sigma_1^2)$ = random effect of membership in group k within condition l

$\gamma_{2, jk:l} \sim N(0, \sigma_2^2)$ = random effect of membership in group k within condition l at time j

$\gamma_{3, i:k:l} \sim N(0, \sigma_3^2)$ = random effect of member

Table 5. Correlations from empirical studies of behavioral HIV prevention for men who have sex with men

Study	Location	$2g$	$\hat{ICC}$	$\hat{r}_{yy(g)}$	$\hat{r}_{yy(m)}$
Cross-sectional					
Kelly 1991	Southern US	3	--	--	--
Kelly 1997	Northern US	8	0.005	0.43	--
Miller 1998	New York City	3	--	--	--
ACDP 1999	Western US	4	0.013	0.71	--
Elford 2002	London	5	0.000	0.59	--
Flowers 2002	Scotland	2	--	--	--
Cohort					
Kegeles 1996	Western US	2	--	--	0.51
Hoff 1997	Western US	2	--	--	0.47
Shepherd 1997	England	2	--	--	0.67
Kegeles 2002	Southwest US	3	--	--	*
Amirkhanian 2005	Russia, Bulgaria	52	0.05	0.02	0.39

$2g$ = number of groups intervention condition plus number of groups in comparison condition; $\hat{ICC}$ = intraclass correlation; $\hat{r}_{yy(g)}$ = over-time correlation at the group level; $\hat{r}_{yy(m)}$ = over-time correlation at the member level; -- not applicable; * not available

Table 6. Estimation of covariance parameters for multiplicative models

Component	Estimation
$\hat{\sigma}^2_g$	$\hat{\sigma}^2_{g,unscaled} \cong \hat{\sigma}^2_{g,scaled} \left(\overline{p}(1-\overline{p})\right)^2$
$\hat{\sigma}^2_{tg}$	$\hat{\sigma}^2_{tg,unscaled} \cong \hat{\sigma}^2_{tg,scaled} \left(\overline{p}(1-\overline{p})\right)^2$
$\hat{\sigma}^2_m$	$\hat{\sigma}^2_{m,unscaled} \cong \hat{\sigma}^2_{m,scaled} \overline{p}(1-\overline{p})$
$\hat{\sigma}^2_e$	$\hat{\sigma}^2_{residual} \cong EDS \times \hat{\sigma}^2_{theoretical} = EDS \times \overline{p}(1-\overline{p})$

Note: $\hat{\sigma}^2_g$ = group component of variance; $\hat{\sigma}^2_{tg}$ = time by group component of variance; $\hat{\sigma}^2_m$ = member component of variance; $\hat{\sigma}^2_e$ = error component of variance; N/A= not applicable; EDS = extra-dispersion scale. Adapted from Murray pages 239 and 306

Table 7. Average results of logistic models applied to simulated group-randomized data in 3 designs, with 2 methods of naïve analysis for the cohort design. Each iteration included 3 groups in each of 2 conditions.

			Design	
	Posttest-only	Cross-sectional	Cohort partly naive	Cohort completely naive
# iterations attempted	1400	1100	3900	3900
# iterations used	960	539	1067	1449
Members per group	300	300	100	100
Total members	1800	1800	600	600
Average fixed parameter estimates				
β0	-1.00	-0.98	-0.60	-0.61
β1		0.00	-0.01	-0.01
β2		0.00	0.00	0.00
β3	-0.34	-0.35	-0.22	-0.22
Average covariance parameter estimates				
M			0.36	0.36
tg:c		0.05	0.04	0.04
g:c	0.06	0.06	0.14	0.14
Err	0.99	0.99	0.62	0.62
Average correlations				
rM			0.37	0.37
rG		0.44	0.46	0.46
ICC linear	0.0113	0.0092	0.0125	0.0129
ICC log1	0.0112	0.0088	0.0124	0.0128
ICC log4	0.0111	0.0087	0.0124	0.0128
ICC logX	0.0109	0.0086	0.0122	0.0126
Average VIF and CF				
VIF linear	4.38	3.7381	2.237	1.447*
VIF log1	4.35	3.6206	2.230	1.439*
VIF log4	4.32	3.5995	2.226	1.437*
VIF logX	4.26	3.5583	2.212	1.427*
CF linear	4.43	3.7391	2.209	1.432
CF log1	4.40	3.6205	2.203	1.424
CF log4	4.37	3.5993	2.199	1.421
CF logX	4.30	3.5580	2.187	1.412
Average estimates of weight				
True	33.63	18.13	15.16	14.97
Naïve	78.71	41.50	26.99	16.34
VIF linear	33.38	17.52	15.06	14.63
VIF log1	33.41	17.84	15.12	14.70
VIF log4	33.50	17.89	15.13	14.71
VIF logX	33.58	17.96	15.18	14.76
CF linear	33.33	17.63	15.32	14.86

CF log1	33.36	17.96	15.39	14.94
CF log4	33.45	18.01	15.40	14.95
CF logX	33.53	18.08	15.44	14.99
MSE of weight				
Naïve	2647	670.7	180.3	46.4
VIF linear	3.47	9.28	1.970	2.570
VIF log1	0.21	0.21	0.019	0.415
VIF log4	0.09	0.17	0.014	0.403
VIF logX	0.08	0.12	0.016	0.349
CF linear	3.49	8.92	2.336	2.176
CF log1	0.24	0.06	0.463	0.018
CF log4	0.11	0.03	0.466	0.014
CF logX	0.09	0.02	0.518	0.015

* For the completely naïve analysis of the nested cohort design, each VIF for each iteration has been multiplied by a factor of $\left(1 - \hat{r}_{yy(m)}\right)$.

Table 8. Impact of ICC correction on weights of empirical studies

Study	2n	2g	m	VIF		Weight				
				Assuming ICC =		Assuming ICC =			Study-specific ICC	Complete analysis
				.005	.03	0	.005	.03		
Flowers 2002	2324	2	1162	6.8	35.8	59.7	8.8	1.7	--	--
Hoff 1997*	1973	2	987	5.9	30.6	59.7	10.1	2.0	--	--
ACDP 1999	1205	4	301	2.5	10.0	12.3	4.9	1.2	2.5	5.8
Kelly 1991	634	3	211	2.1	7.3	15.2	7.4	2.1	--	--
Kegeles 2002*	632	3	211	2.0	7.3	15.2	7.4	2.1	--	--
Elford 2002	1011	5	202	2.0	7.0	23.2	11.6	3.3	23.2	23.2
Miller 1998	385	3	128	1.6	4.8	8.2	5.0	1.7	--	--
Kegeles 1996*	188	2	94	1.5	3.8	11.0	7.5	2.9	--	--
Kelly 1997	385	8	48	1.2	2.4	8.7	7.1	3.7	7.1	8.3
Shepherd 1997*	54	2	27	1.1	1.8	2.5	2.2	1.4	--	--
Amirkhanian 2005*	210	52	4	1.0	1.1	9.6	9.4	8.8	8.3	12.3

* Nested cohort design. Unmarked studies used the nested cross-sectional design

2n for cohort studies = sample size at followup including both conditions
2n for cross-sectional studies = average of baseline and followup sample sizes including both conditions
2g = total number of groups in the two conditions
m = n / g
VIF = 1+ (m-1) × ICC for nested cross-sectional studies;
VIF = [1+ (m-1) × ICC] $(1 - \hat{r}_{yy(m)})$ for nested cohort studies

Complete analysis = Mixed logistic model of raw data assuming ICC as estimated from the data and accounting for design features such as multiple waves of followup data

Table 9. Meta-analyses of community-level HIV prevention interventions for men who have sex with men

Assumptions	Fixed effects					Random effects	
	Total weight	OR (95% CI)	Q (p)	*df*	tau	Total weight	OR (95% CI)
ICC=0	225.3	.75 (.66, .85)	14.5 (.15)	10	.02	128.7	.71 (.60, .85)
ICC=0.005	81.5	.67 (.54, .84)	8.6 (.57)	10	.00	81.5	.67 (.54, .84)
ICC=0.03	30.8	.58 (.41, .83)	4.0 (.95)	10	.00	30.8	.58 (.41, .83)
Study-specific ICC*	89.6	.73 (.59, .90)	11.2 (.35)	10	.01	77.1	.71 (.57, .89)
Complete analyses**	98.1	.72 (.59, .87)	10.6 (.39)	10	.01	90.6	.71 (.58, .87)
Only the 4 complete analyses	49.6	.74 (.56, .97)	9.4 (.02)	3	.44	13.9	.65 (.38, 1.09)

* Assuming ICC = .005 if ICC not available from study
** Complete analyses for ACDP 1999, Amirkhanian 2005, Elford 2002, and Kelly 1997; otherwise same as analyses based on study-specific ICC

Appendix: CORRECTION FACTORS FOR NESTED CROSS-SECTIONAL AND

NESTED COHORT DESIGNS

The Nested Cross-Sectional Design

In the nested cross-sectional group-randomized trial design (Murray pages 140-143), each group (e.g., school, hospital, neighborhood, etc) is randomly assigned either to the experimental or the control condition. Pretest and posttest measurements are taken among cross-sectional samples of members of each group, but no effort is made to obtain responses from the same individuals at posttest who were surveyed at pretest. The model is described by Murray for the linear situation as follows (pg 140-141):

$$Y_{i:jk:l} = \mu + C_l + T_j + TC_{jl} + \boldsymbol{G}_{k:l} + \boldsymbol{TG}_{jk:l} + \boldsymbol{\varepsilon}_{i:jk:l}$$

[$\boldsymbol{G}$, $\boldsymbol{TG}$, and $\boldsymbol{\varepsilon}$ are random effects (in bold)]

The observed value $Y_{i:jk:l}$ for the ith member nested within the kth group and l^{th} condition and observed at the jth time is expressed as a function of a grand mean μ, the effect of the lth condition (C_l), the effect of the jth time (T_j), the joint effect of the lth condition and the jth time (TC_{jl}), the realized value of the kth group ($\boldsymbol{G}_{k:l}$), and the realized value of the combination of the kth group and j^{th} time ($\boldsymbol{TG}_{jk:l}$). Any difference between this predicted value and the observed value is allocated to the residual error ($\boldsymbol{\varepsilon}_{i:jk:l}$).

In most group-randomized trials, condition, time, and their interaction are fixed effects. In order to account for the positive intraclass correlation expected in the data, $\boldsymbol{G}_{k:l}$ and $\boldsymbol{TG}_{jk:l}$ must be included in the analysis as random effects. The three random effects allow for correlation among members within a group ($\boldsymbol{G}_{k:l}$), for correlation among members within a time × group survey ($\boldsymbol{TG}_{jk:l}$) and for random variation among the members ($\boldsymbol{\varepsilon}_{i:jk:l}$).

Similar procedures to those shown for the posttest-only design can be used to

show that the correction factor for the nested cross-sectional design is

$$cf_{cross-sectional} = \frac{(n-1)(m-1)VIF_{m:tg:c}(1-r_{yy(g)})}{(m-1)(n-VIF_{m:tg:c})-(n-1)(m-VIF_{m:tg:c})r_{yy(g)}} \qquad 5$$

where

$$VIF_{m:tg:c} = 1 + (m-1)ICC_{m:tg:c}$$

$$ICC_{m:tg:c} = \frac{\sigma_{tg:c}^2}{\sigma_{tg:c}^2 + \sigma_e^2}$$

$$r_{yy(g)} = \frac{\sigma_{g:c}^2}{\sigma_{g:c}^2 + \sigma_{tg:c}^2}$$

If $ICC_{m:tg:c}$ is small, then $n - VIF_{m:tg:c}$ and $m - VIF_{m:tg:c}$ approximately equal $n-1$

and $m-1$. The factor $1 - r_{yy(g)}$ cancels out and the correction factor is approximately the

VIF.

The Nested Cohort Design

In the nested cohort group-randomized trial design (Murray pg 179-184, 370),

each group is randomly assigned either to the experimental or the control condition.

Pretest and posttest measurements are taken among a cohort of members of each group,

with responses from the same individuals at posttest who were surveyed at pretest. We

adapt the model described by Murray (pg 180-181) as follows:

$$Y_{ij:k:l} = \mu + C_l + T_j + TC_{jl} + \mathbf{G}_{k:l} + \mathbf{M}_{i:k:l} + \mathbf{TG}_{jk:l} + \mathbf{\varepsilon}_{ij:k:l}$$

$[\mathbf{G}, \mathbf{M}, \mathbf{TG},$ and ε are random effects (in bold)]

The observed value $Y_{ij:k:l}$ for the ith member at the jth time and nested within the kth

group and l^{th} condition is expressed as a function of a grand mean μ, the effect of the lth

condition (C_l), the effect of the jth time (T_j), the joint effect of the jth time and the lth condition (TC_{jl}), the realized value of the kth group ($G_{k:l}$), the realized value of the ith member ($M_{i:k:l}$), and the realized value of the combination of the jth time and k^{th} group ($TG_{jk:l}$). Any difference between this predicted value and the observed value is allocated to the residual error ($\varepsilon_{ij:k:l}$).

In most group-randomized trials, condition, time, and their interaction are fixed effects. In order to account for the positive intraclass correlation expected in the data, $G_{k:l}$ and $TG_{jk:l}$ must be included in the analysis as random effects. The three random effects carried over from the nested cross-sectional design allow for correlation among members within a group ($G_{k:l}$), for correlation among members within a time × group survey ($TG_{jk:l}$) and for random variation among the members ($\varepsilon_{ij:k:l}$).

Murray includes an additional term $MT_{ij:k:l}$ to allow for correlation among replicate measurements on the same member during a single time during the survey. We assume no replicate measurements and therefore exclude this term. Procedures similar to those shown for the posttest-only design can be used to show that the correction factor for the nested cohort design is

$$cf_{cohort} = \frac{(n-1)(m-1)VIF_{mt:g:c}\left(1-r_{yy(g)}\right)\left(1-r_{yy(m)}\right)}{(m-1)\left(n-VIF_{mt:g:c}\right)-(n-1)\left(m-VIF_{mt:g:c}\right)r_{yy(g)}-m(g-1)\left(VIF_{mt:g:c}-1\right)r_{yy(m)}} \qquad 6$$

where

$$VIF_{mt:g:c} = 1+(m-1)ICC_{mt:g:c}$$

$$ICC_{mt:g:c} = \frac{\sigma^2_{tg:c}}{\sigma^2_{tg:c}+\sigma^2_e}$$

$$r_{yy(g)} = \frac{\sigma^2_{g:c}}{\sigma^2_{g:c}+\sigma^2_{tg:c}}$$

$$r_{yy(m)} = \frac{\sigma^2_{m:g:c}}{\sigma^2_{m:g:c} + \sigma^2_e}$$

Murray includes an additional term $\sigma^2_{mt:g:c}$ in the denominators of $r_{yy(m)}$

and $ICC_{mt:g:c}$ (pages 300-301), but because we do not allow for replicate observations on

an individual during a single time point in the survey our denominators do not include

this term.

When the individual data are not available and only a summary two-by-two

contingency table is provided without regard to group membership, the ICC and pre-to-

post correlations (or the corresponding covariance parameters) will not usually be

available. Therefore it is necessary to borrow estimates of ICC and pre-to-post

correlations from similar studies in order to apply this procedure.

One cluster per condition

Because we rely on an external estimate of ICC, these correction factors can be

used even when only one group is assigned to intervention and one to control status, that

is, when $g = 1$. In this case, $n = m$ and the corrected variance for both the posttest-only

design and the nested cross-sectional design further simplifies to

$$\sigma^2_{\Delta, A} = \frac{VIF}{1 - ICC} \sigma^2_{\Delta, U}$$

The corrected variance for the nested cohort design simplifies to

$$\sigma^2_{\Delta, A} = \frac{VIF\left(1 - r_{yy(m)}\right)}{1 - ICC} \sigma^2_{\Delta, U}$$

This formula corrects only the variance in such studies, and cannot address bias in

the point estimate of the intervention effect that is likely when only one cluster is

assigned to each treatment condition. In the context of meta-analysis however, that bias is

diminished when several studies are aggregated, if there is no publication bias.

Appendix B. Transformations Between Count And Dichotomous Outcome Measures For Meta-Analysis:

Estimating Risk Ratios from Means and Rate Ratios from Proportions

TRANSFORMATIONS BETWEEN COUNT

AND DICHOTOMOUS OUTCOME MEASURES FOR META-ANALYSIS

Estimating Risk Ratios from Means and Rate Ratios from Proportions

DRAFT DATE　　　　July 30, 2009

Word count:　　　　　7847 (main text only, excluding appendices, tables, refs, etc.)

Running header:　　　Transformations Between Count and Dichotomous Data

Wayne D. Johnson

e-mail: wjohnson@cdc.gov

ABSTRACT

Meta-analysis requires comparable units or metrics (e.g., risk ratios, rate ratios, odds ratios, or standardized mean differences) to combine and compare information across studies. When the outcome of interest is a count variable that can be dichotomized, some studies may provide only means and variances (or standard deviations), while others treat the outcome as dichotomous and provide only proportions. The count variables are often highly skewed and not normally distributed. We used regression models and the method of moments to identify candidate procedures for estimation of proportions (and therefore risk ratios) from means and variances, and estimation of means and variances (and therefore rate ratios) from proportions. We used these methods as well as the delta method to estimate variances of the intervention effects. We compared performance of the various methods on empirical studies when the proportions, means, and variances were available. The empirical data were not normally distributed. The most accurate estimates of risk ratios and their variances from means were obtained using regression estimates or the method of moments assuming a negative binomial distribution. There was little difference among several approaches for estimating rate ratios from dichotomous data, but the variance of the log rate ratio was best estimated again by way of the regression estimates or the method of moments assuming a negative binomial distribution. With the caveat that intervention effects on dichotomous and count (or normally distributed) outcomes are not precisely interchangeable, these formulas facilitate comparison and aggregation for meta-analysis and health policy.

INTRODUCTION

Meta-analysis requires comparable metrics (e.g., risk ratios, rate ratios, odds ratios, or standardized mean differences) to combine and compare information across studies. But research protocols differ among studies, and various outcome measures are likely to be encountered in a given field. For example, in HIV prevention research, unprotected sex may be measured as a dichotomous variable (any *vs.* none) or as a count variable (number of occasions of or partners for unprotected sex). The intervention effect might then be measured as the *risk ratio* comparing the proportion reporting one or more occasions, or as the *rate ratio* comparing the mean number of events per unit time during the recall period in the intervention versus the comparison condition (Table 1). A method to transform results between dichotomous and count measures is needed to permit comprehensive analysis of randomized controlled trials which use the two types of outcomes.

---- Table 1 about here ---

A method has previously been described for transformation between intervention effects measured as odds ratios or standardized mean differences as defined in Table 1 ([1],[2],[3]). Although valid odds ratios and standardized mean differences can be estimated for randomized trials, there are several reasons why risk ratios and rate ratios may be preferred. First, the risk ratio and rate ratio are directly interpretable as the ratio of either the proportion reporting any occasions or the mean number of occasions per unit time between the two treatment conditions. Second, odds ratios can be misleading: when events are common, the odds ratio can be much farther from the null than the risk ratio, and the distinction between the two is frequently missed ([4],[5],[6],[7],[8],[9]). Third, in a series of

studies we are interested in, the outcome is a count variable that is not normally distributed; the empirical performance of the method that is currently being applied has not been evaluated for this situation. However, in the absence of methods to transform directly between risk ratios and rate ratios (and their variances), meta-analyses of randomized controlled trials have had to rely on transformations between odds ratios and standardized mean differences ([3]).

Here we derive and examine the performance of alternative methods to transform data between risk ratios for dichotomies and rate ratios for counts. We consider methods based on linear associations between count and dichotomous measures across studies, as well as the method of moments ([10]). We derive formulas to estimate proportions given only means and variances, and then to estimate means and variances given only proportions. We then substitute these estimates into the usual formulas to estimate intervention effects and study weights (the reciprocal of the variance of the intervention effect). We also consider the delta method for estimation of study weights ([11]). We apply these methods to the empirical studies and compare the success of the various approaches in terms of mean squared error (MSE) compared to observed estimates of intervention effects and weights.

METHODS

For this analysis we selected studies for which means, variances, proportions, and sample sizes were available for intervention and comparison groups at any number of time points, typically one baseline and one or more follow up times such as 6 months and 12 months. Search strategies have been described previously ([12]). We identified 19 studies of HIV prevention for which both dichotomous and count outcomes were

available ([13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [23] [24] [25] [26] [27] [28] [29] [30] [31]). Two studies provided a total of 8 independent intervention/comparison pairs for a total of 25 pairs. Most studies included multiple time points for a total of 79 comparisons of intervention and control data.

Risk ratios and rate ratios

Here we provide the well-known expressions for the risk ratio and variance of the log risk ratio (VarLnRRisk) given proportions and sample sizes for each treatment condition. Similarly, we give expressions for the rate ratio and variance of the log rate ratio (VarLnRRate) given means, variances, and sample sizes. These expressions for VarLnRRisk and VarLnRRate arise from the delta method with the assumption of independence between comparison and intervention groups ([11], and see Appendix 1). We then describe several candidate methods for estimating risk ratios and VarLnRRisk given only means, variances, and sample sizes, and for estimating rate ratios and VarLnRRate given only proportions and sample sizes.

We define the risk ratio as the ratio of the average risk p_1 of one or more events among members of the intervention group during a specified recall period to the average risk p_2 of one or more events during the same recall period among members of the comparison group. The risk ratio and VarLnRRisk are routinely calculated from the sample sizes n_1 and n_2 and observed sample proportions $\hat{p}_1$ and $\hat{p}_2$ for the intervention and control groups as follows:

$$\hat{Risk}\ Ratio = \frac{\hat{p}_1}{\hat{p}_2} \qquad \qquad \text{Equation (1)}$$

$$\hat{Var}\left(LnR\hat{Risk}\right) = \frac{1-\hat{p}_1}{n_1\hat{p}_1} + \frac{1-\hat{p}_2}{n_2\hat{p}_2} \qquad \qquad \text{Equation (2)}$$

Similarly, we define the rate ratio as the ratio of the average number of events μ_1 per unit time during a specified recall period among members of the intervention group to the average number of events μ_2 per unit time during the same recall period among members of the comparison group. The rate ratio and VarLnRRate are estimated from the sample sizes and the observed sample means $\hat{\mu}_1$ and $\hat{\mu}_2$ and sample variances $V\hat{a}r\ Y_1$ and $V\hat{a}r\ Y_2$ for the count variable Y_1 in the intervention group and Y_2 in the control group:

$$R\hat{a}te\ Ratio = \frac{\hat{\mu}_1}{\hat{\mu}_2} \qquad\qquad \text{Equation (3)}$$

$$V\hat{a}r\left(LnR\hat{R}ate\right) = \frac{V\hat{a}r\ Y_1}{n_1\hat{\mu}_1^2} + \frac{V\hat{a}r\ Y_2}{n_2\hat{\mu}_2^2} \qquad\qquad \text{Equation (4)}$$

In the scenarios of interest, the sample sizes and limited statistics are given. We may be given only means and variances (or standard deviations) of the number of occasions during a specified recall period. In that case we need to estimate proportions reporting any occasions during the recall period and to use those estimated proportions to estimate the risk ratio.

Alternatively, we may be given proportions reporting any occasions during the recall period and need to estimate means and variances of the number of occasions during the recall period. Those estimated means are then used to estimate the rate ratio.

For transforming in each direction, we considered five candidate methods for estimating risk ratios and rate ratios. A first naïve estimate was obtained by simply interchanging the risk ratio and rate ratio as if they were equivalent; we refer to this as the identity method. A second set of estimates was obtained empirically by linear regression, relating the logit of the observed proportion with the corresponding log of the mean and

the log of the variance, across studies. Third, fourth, and fifth sets of estimates were

obtained by way of the method of moments, assuming either a Poisson, geometric, or

negative binomial distribution of the count variable.

Precision-based weights for each study, as is common practice ([32]), are calculated

for meta-analysis as the reciprocal of VarLnRRisk or VarLnRRate. We examined

methods for estimating variances of the intervention effects that correspond with each

method for estimating risk ratios or rate ratios by substituting the estimates of proportion

given only mean and variance, or vice versa, into the variance formulas (2) and (4) above.

Thus the sample sizes n_1 and n_2 and estimated proportions $\hat{p}_1$ and $\hat{p}_2$ (as estimated by

means and variances) for the intervention and control groups were substituted into

equations (1) and (2) to estimate the risk ratios and VarLnRRisk. Similarly, the sample

sizes and the estimated means $\hat{\mu}_1$ and $\hat{\mu}_2$ and estimated variances $V\hat{a}r\ Y_1$ and $V\hat{a}r\ Y_2$ (as

estimated by proportions) for the intervention and control groups were substituted into

equations (3) and (4) to estimate the rate ratios and VarLnRRate. We refer to this as the

substitution method for estimating the variance (and its reciprocal the weight) of the

intervention effect.

Although it will be seen that some of these approaches provide results that very

nearly approximate the values reported or calculated from the actual data for the

variances and weights, it must be emphasized that the quantities resulting from this

substitution process are not actually statistical variances, in the sense of average squared

distance from the mean, of the log risk ratio (LnRRisk) given only means and standard

deviations, or of the log rate ratio (LnRRate) given only proportions. Instead, they

represent an attempt to estimate the variances that would have been observed for the

LnRRisk if the proportions had been reported, or for the LnRRate if the means and standard deviations had been reported. To address this limitation, we also estimated the variance for each intervention effect by way of the delta method. Each approach for estimating LnRRisk, LnRRate, and corresponding variances is described in detail in the following paragraphs.

Interchanging risk ratios and rate ratios

The simplest approach to estimating risk ratios given only means, or to estimating rate ratios given only proportions, is to substitute rate ratio for risk ratio and vice versa. We will refer to this as the Identity approach. This approach could reflect the interpretation of a reader who is familiar with the concept of ratios as a measure of intervention effects but who may not attend to the more subtle differences among risk ratios and rate ratios. While it might seem intuitively likely that proportions would be correlated with counts and that risk ratios would be correlated with rate ratios, there is no statistical relationship stipulating that one quantity must necessarily increase with the other, and, to our knowledge, the strength of these correlations has not yet been examined empirically. For this first naïve procedure we also interchanged the variance (and weight) of the LnRRisk with the variance (and weight) of the LnRRate.

Linear regressions

We based the next approach on empiric associations between the logit of the proportion reporting one or more occasions and the corresponding log of the mean number of occasions within that study as well as the log of the variance in number of occasions evaluated across studies. The pattern turned out to be approximately linear (on a log-log scale) across studies (Figure 1). Because all three of these parameters range

conceptually from negative infinity to positive infinity, this approach could represent an improvement over interchanging risk ratios and rate ratios where proportions are bounded at both ends of the range (from 0% to 100%) while counts are bounded only at the lower end (from 0 to infinity). The regression approach is also supported by the empirically observed relationship between means and proportions in the studies which provide both.

We then applied the substitution method as described above to estimate variances of the intervention effects. That is, the sample sizes n_1 and n_2 and regression-based estimates of proportion $\hat{p}_1$ and $\hat{p}_2$ (as estimated by means and variances) for the intervention and control groups were substituted into equations (1) and (2) to estimate the risk ratios and VarLnRRisk. Similarly, the sample sizes and the regression-based estimates of mean $\hat{\mu}_1$ and $\hat{\mu}_2$ and variance $V\hat{a}r\ Y_1$ and $V\hat{a}r\ Y_2$ (as estimated by proportions) for the intervention and control groups were substituted into equations (3) and (4) to estimate the rate ratios and VarLnRRate. The weights in either case (risk ratio or rate ratio) were obtained as the reciprocal of the variance.

The regression-based approach is useful if the correlations among means, variances, and proportions are strong and consistent, and if a sufficient number of studies are available to estimate the linear associations. Both of these criteria are met in our empirical example, as shown below.

Method of moments

The next approach does not require regression or a large number of empirical studies to estimate rate ratios from risks and conversely. Instead, the method of moments allows estimation of the proportion given the mean and variance, or vice versa, by assuming some specific underlying parametric distribution of the data. First, we derive

the appropriate formulas under each candidate distribution. We present results from two 1-parameter distributions (Poisson and geometric) and one 2-parameter distribution (the negative binomial distribution). Once estimates of proportion given mean and variance, , are obtained by way of the method of moments, we proceed again as before to substitute the estimates into Equations 1 and 2 for the estimate and variance of the risk ratio. Similarly we substitute estimates of mean and variance given proportion into Equations 3 and 4 for the estimate and variance of the rate ratio. Again, we emphasize that this substitution method does not yield a *bona fide* estimate of the variance, as it does not take into account the uncertainty inherent in the estimation of the proportion given sample mean and variance, or vice versa. Instead, this method estimates a value for the variance that we might have calculated if we had directly observed the relevant statistics (the sample proportion for risk ratio, or the sample mean and variance for rate ratio). Thus, it cannot be interpreted as a true measure of average dispersion from the mean.

Derivations for transforming means and variances to proportions

Expressions for the mean, variance, and probability function (pf) are given for each of the three candidate distributions (Table 2). The proportion p, which is the proportion where the outcome value is greater than zero, can then be expressed as $1 - P[Y=0]$. We then express the various parameters (μ, g, or k) in terms of the mean and variance under each distribution (Appendix 2). The estimates of proportion obtained by replacing the population mean and variance by their sample estimates can then be substituted into Equations (1) and (2) for risk ratios and VarLnRRisk (Appendix 3). Again, this substitution approach does not yield a statistically interpretable estimate of VarLnRRisk.

---- Table 2 about here ---

Derivations for transforming proportions to means and variances

Next, we examined transformations in the reverse direction, where we use the observed proportion to estimate the mean and variance. Because we had only one summary statistic (the sample proportion), we were restricted to modeling the underlying count data by a 1-parameter distribution. We used the relationships from Table 2 for the Poisson and geometric distributions, and estimated the mean and variance in terms of the observed sample proportion $\hat{p}$ (Appendices 4 & 5, Table 3).

---- Table 3 about here ---

Because we are given only one summary statistic (the proportion), some type of simplification of the two parameters of the negative binomial distribution was required in order to simultaneously estimate both the mean and variance. As shown above, we found a strong linear association between the log of the sample variance $\ln S^2$ and the log of the sample mean $\ln \overline{Y}$ in each study condition at each time point. Estimating this relationship via linear regression, we were able to parameterize fully the negative binomial distribution in terms of the population proportion and thus estimate the unobserved mean and variance as known functions of the proportion estimate. This allows us to estimate the unobserved mean by equating the expressions for proportion in table 2 to the observed proportion. With this assumption, the two parameters of the negative binomial distribution can be reduced to one parameter. The resulting estimates for the mean have no closed algebraic form. They are expressed implicitly in table 3 and can be solved numerically for each study.

Delta method

Finally, the delta method applies only to estimation of variances, not the point estimates of LnRRisk and LnRRate. In general, the delta method approximates the expectation of a function by the expected value of an approximation to the function ([11]). The typical application is to estimate the squared error from the mean and accordingly to estimate the variance. Thus, the delta method addresses the lack of statistical justification for the substitution method.

Application to empirical examples

Next we used empirical data from studies of HIV prevention interventions to compare the performance of the resulting candidate formulas. For each of the five approaches (interchanging risk ratios and rate ratios; regression; or method of moments assuming Poisson, geometric, or negative binomial distributions), we calculated the mean squared error (MSE) of the intervention effect as the average square of the difference between the intervention effects (LnRRisk or LnRRate) as estimated by each method versus the observed intervention effects. Similarly, we calculated the MSE of the corresponding weight estimates as the average square of the difference between the natural log of the estimated weight of the study by each method versus the natural log of the observed weight of the study. We then compared the MSEs from the various methods to determine which approaches yielded the most accurate estimates of intervention effects (either LnRRisk or LnRRate) and weights (the reciprocal of either VarLnRRisk or VarLnRRate). Thus,

$$MSE = \frac{\sum_{i=1}^{79}(P_i - E_i)^2}{79}$$

where

E_i = estimated value of LnRRisk, LnRRate, or log of weight

O_i = observed value of LnRRisk, LnRRate, or log of weight

for each of 79 combinations of study by time point.

Because intervention effects from a given study at different times are not

independent, and because intervention effects at baseline cluster around the null value,

we repeated this process using results from only a single follow-up time point from each

study or subsection of each study, resulting in 25 observations. The results were similar.

Data from this subset analysis appears in two figures (3 and 5) so the data points are more

easily distinguishable.

RESULTS

The empirical count data that we found do not approximate a normal distribution.

Figure 1 shows the baseline distribution for intervention and control conditions combined

for several studies when the raw count data were available. To reduce clutter in the

figure, we truncated the data plots for the eight separate studies at 19 episodes of UAI.

When the data from these studies are combined, the counts decrease in a remarkably

consistent linear trend on this log-log scale. We used running averages to smooth the

trend as counts become sparse toward the right side of the figure. Overall, about 66% of

respondents reported 0 occasions of unprotected sex, 12% reported 1 occasion, and 6%

reported 2 occasions. Only 10% reported 5 or more occasions, 5% reported more than 10

occasions, and 1% reported more than 50 occasions. The data might be better fit by some

other parametric distribution, such as a zeta or an approximate gamma distribution, when

the number of respondents decreases consistently as the value of the count variable

increases; we address this issue in the discussion ([33]).

---- Figure 1 about here ---

Regression parameters from empirical studies

There were strong correlations among the observed 158 means, variances, and proportions. The logs of the means predicted 92% of the variance among the logs of the variances and 60% of the variance among the logits of the proportions (Figure 2).

---- Figure 2 about here ---

We summarize the various relationships between the logit proportion (logit p), the population mean (E[Y]), and the population variance (Var [Y]), as follows, where the quantities on the left sides are to be estimated in terms of those on the right sides.

$$\text{logit}(p) = 1.63 \times \ln E[Y] - 0.57 \times \ln Var[Y] + 0.04 \qquad \text{Equation (5)}$$

$$\ln E[Y] = 1.13 \times \text{logit}(p) + 1.76 \qquad \text{Equation (6)}$$

$$\ln Var[Y] = 1.80 \times \text{logit}(p) + 5.34 \qquad \text{Equation (7)}$$

$$\ln Var[Y] = 1.93 \times \ln E[Y] + 2.17 \qquad \text{Equation (8)}$$

We assume these relationships do not differ between intervention and comparison or across time. The R-squared values (the proportion of variance among the values on the left hand side explained by the values on the right hand side) for equation (5) is .81, for (6) is .60, for (7) is .38, and for (8) is .92. Equation (8) implies that the variance of the counts in each study arm is approximately a function of the square of the mean:

$$Var[Y] = e^{2.17} E[Y]^{1.93} = 8.76 E[Y]^{1.93} \qquad \text{Equation (9)}$$

Estimates of proportion given mean and variance from equation 5 can then be substituted into equations 1 and 2 for risk ratios and VarLnRRisk for comparison to the directly observed values. Similarly, estimates of mean and variance given proportion from equations 6 and 7 can be substituted into equations 3 and 4 for rate ratios and

VarLnRRate. Again, these numerical estimates are not genuine variances in the sense of statistical deviation from the mean.

To obtain regression parameters for the analyses below, we excluded one study at a time in 19 separate iterations of each the regressions represented above by equations 5 through 8. For each study we used regression parameters obtained only from the other 18 studies, thus simulating a situation in which parameters for studies which provide only one type of data (either count or dichotomous) must be estimated from other studies which provide both. The resulting four series of 19 regressions differed little from the equations shown above. For example, the coefficient of $\ln E[Y]$ in the 19 iterations of equation (8) was between 1.90 and 1.96 in all but one case (1.99). The constant was between 2.13 and 2.24 in all but one case (2.02). These two exceptions were for different studies. R-squared values for those 19 models ranged from 0.91 to 0.94.

Application of the transformation methods to empirical studies

Estimating the risk ratio and weight (reciprocal of VarLnRRisk) given only means, standard deviations, and sample sizes

The MSE between LnRRisk and LnRRate was 0.34, with only a small absolute bias of 0.02 (Table 4). But MSE between log of study weights for dichotomous versus count outcomes was 2.58. The average bias in log weights of 1.47 indicates that study weights for dichotomous outcomes were on average $e^{1.47} = 4.35$ times the weights for the same studies when outcomes were measured by count variables.

---- Table 4 about here ---

When comparing values of LnRRisk estimated by various methods to the observed LnRRisk, the regression model estimating the logit of p as a joint function of

the log of the mean and the log of the variance performed best in that it yielded the least

MSE (or the greatest precision) among the options we considered (MSE = 0.05). The

same model yielded the least MSE in estimating study weights (MSE = 0.10). Regression

also yielded only a small absolute bias for the LnRRisk (bias = 0.03) and study weights

(bias =-0.0011). Among the models that relied on the method of moments, the closest

correspondence between estimated and observed values was obtained under the

assumption of the negative binomial distribution, where the MSE was 0.09 for the

LnRRisk and 0.15 for study weights; bias was also minimal under the method of

moments with the negative binomial assumption.

     To illustrate these results, we compared estimates of LnRRisk by several methods

with only means and variances provided rather than proportions (Figure 3, note

logarithmic scale; LnRRisk for two studies fall outside the range shown). On the

horizontal axis, we reported the observed LnRRisk and on the vertical axis, we reported

the observed LnRRate and two different estimates of the LnRRisk based on the observed

means and variances. Taking the study at the far left as an example, the observed

LnRRisk of -0.89 indicates a risk ratio of 0.41, or 59% fewer individuals reporting any

unprotected sex in the intervention group than in the comparison group. The LnRRate for

this study of -0.36 indicates a rate ratio of 0.70, or only 30% fewer episodes of or partners

for unprotected sex in the intervention group than in the comparison group. The estimated

LnRRisk of -0.69 obtained by way of regression estimates indicates a risk ratio of 0.52,

or 48% fewer individuals reporting any unprotected sex in the intervention group than in

comparison. Finally, the estimated LnRRisk of -0.87 obtained by way of the negative

binomial distribution indicates a rate ratio of 0.42, or 58% fewer individuals reporting

any unprotected sex in the intervention condition than in the comparison condition. Although the negative binomial estimate in this study was closer than the regression estimate to the observed value, on average, the best estimates of LnRRisk given only means and variances were obtained by way of the regression estimates followed by the negative binomial, and both performed substantially better than simple substitution of LnRRate in place of LnRRisk.

--- Figure 3 about here ---

We can also compare estimated weights for each LnRRisk by the same methods (Figure 4). Substitution of the weight of the LnRRate yields estimates that are mostly much smaller than the observed weight of the LnRRisk. For example, the observed weight of the LnRRisk for the study at the far right of the graph is 2392 ($\approx e^{7.8}$), while the observed weight of the LnRRate is only 139 ($\approx e^{4.9}$); the discrepancy is a factor of over 17.

The regression approach yields values that much more closely approximate the observed weights of the LnRRisk. By this method, the weight estimated for the same study is 1125 ($\approx e^{7.0}$), which is off by a factor of 2.1. Estimates of weight based on substituting proportions obtained from the method of moments assuming an underlying negative binomial distribution are also closer to the observed weights. By this approach, the estimated weight of the same study is 880 ($\approx e^{6.8}$), which is off by a factor of 2.7.

--- Figure 4 about here ---

Estimating the rate ratio and weight (reciprocal of VarLnRRate) given only proportions and sample sizes

As has already been observed, the MSE between LnRRate and LnRRisk was 0.34, with only a small absolute bias of 0.02 (Table 5). When estimating LnRRate, essentially the same MSE and bias were observed for each of the four alternative methods (regression or method of moments assuming Poisson, geometric, or negative binomial distributions) as for the naïve first procedure. When estimating study weights, the regression model yielded the least MSE (0.32) and bias (0.06). Among the models which relied on the method of moments, the closest correspondence for study weights was again obtained under the assumption of the negative binomial distribution, when the MSE was 0.38 and bias was 0.07. Thus for the purpose of estimating LnRRate given only proportions, none of these approaches improved on simply substituting the LnRRisk, but the weights for the LnRRate metric were estimated much more accurately by regression or by the method of moments assuming a negative binomial distribution than by assumption of a Poisson or geometric distribution.

--- Table 5 about here ---

To illustrate these results, we compared estimates of LnRRate by three methods with only proportions provided rather than means and variances (Figure 5, note logarithmic scale). In this figure, the negative binomial estimates are very similar to the observed LnRRisk (not the observed LnRRate). The Poisson and geometric estimates (not shown) were always between the observed LnRRisk and the regression estimate of LnRRate, with the regression-based estimate always farthest from the null value of zero on the vertical axis. As an example, the LnRRate of -0.97 observed for the leftmost study indicates a rate ratio of 0.38, or 62% fewer episodes of or partners for unprotected sex in the intervention group than in the comparison group. The LnRRisk for this study of -0.18

indicates a risk ratio of 0.83, or only 17% fewer individuals reporting unprotected sex in the intervention group than in the comparison group. The estimated LnRRate of -0.37 obtained by way of regression estimates indicates a rate ratio of 0.69, or 31% fewer episodes of or partners for unprotected sex in the intervention group than in comparison. Finally the estimated LnRRate of -0.21 obtained by way of the negative binomial distribution indicates a rate ratio of 0.81, or 19% fewer episodes of or partners for unprotected sex in the intervention condition than in the comparison condition; this value is similar to the risk ratio (not the rate ratio) of 0.83.

--- Figure 5 about here ---

We can also compare estimated weights for each LnRRate by the same methods (Figure 6). The pattern is essentially the reverse of that observed for estimated weights of LnRRisk. Substitution of the weight of the LnRRisk yields estimates that are mostly much larger than the observed weight of the LnRRate. The observed weight of the LnRRate for the study at the far right of the graph is 149 ($\approx e^{5.0}$), while the observed weight of the LnRRisk is 1346 ($\approx e^{7.2}$); therefore simple substitution of the LnRRisk weight in a meta-analysis of LnRRate would result in this study being weighed 9 times as heavily as it should be, based on the sample means and variances.

--- Figure 6 about here ---

The regression approach yields values much closer to the approximate observed weights of the LnRRate. By this method, the weight estimated for the same study is 137 ($\approx e^{5.3}$), which is too high by only 37%. Weights estimated by substituting means and variances obtained from the method of moments assuming an underlying negative

binomial distribution are also close to the observed weights. By this approach, the estimated weight of the same study is 282 ($\approx e^{5.6}$), which is off by a factor of about 1.9.

The regression approach to estimating weights for the LnRRate can be simplified (and improved according to one criterion which we explain below) because the exponent of 1.93 (see equation 9) in the numerator of each term of the variance of LnRRate (equation 4) approximately coincides with the exponent of 2 in the denominator. This relationship implies that among the studies at hand, estimates of weight by this method are almost independent of the estimates of means and variances of the individual responses.

$$\hat{Var}\left(LnR\hat{R}ate\right)=\frac{\hat{Var}\ Y_1}{n_1\hat{\mu}_1^2}+\frac{\hat{Var}\ Y_2}{n_2\hat{\mu}_2^2}$$

$$=\frac{e^{2.17}\hat{\mu}_1^{1.93}}{n_1\hat{\mu}_1^2}+\frac{e^{2.17}\hat{\mu}_2^{1.93}}{n_2\hat{\mu}_2^2}=\frac{8.76}{n_1\hat{\mu}_1^{.07}}+\frac{8.76}{n_2\hat{\mu}_2^{.07}} \qquad \text{Equation (10)}$$

Note that $\hat{\mu}^{1.93}/\hat{\mu}^2=\hat{\mu}^{.07}$ is close to 1 across a wide range of values of $\hat{\mu}$. In the studies we identified, the mean ranges from 0.0625 to 13.8, so $\hat{\mu}^{.07}$ ranges from .82 to 1.20. If we evaluate $\hat{\mu}^{.07}$ at the median 3.1 of all $\hat{\mu}$ (regardless of treatment condition) then $\hat{\mu}^{.07}\approx 3.1^{.07}=1.08$ and

$$\hat{Var}\left(LnR\hat{R}ate\right)=\frac{8.76}{n_1\hat{\mu}_1^{.07}}+\frac{8.76}{n_2\hat{\mu}_2^{.07}}\approx\frac{8.1}{n_1}+\frac{8.1}{n_2} \qquad \text{Equation (11)}$$

When we apply this approach in equation (11), the MSE for weight is 0.42 and the absolute bias is 0.07, not quite as precise as the MSE of 0.32 and bias of 0.06 for regression by equation (10) in Table 5. (In both cases we use the series of 19 regressions excluding one study at a time.) The weight estimated for the farthest right data point is 112, or 24% less than the observed weight of 149, but for most of the extreme values,

substitution of the geometric mean as in equation (11) results in more precise estimates than equation (10). Because the previous estimates for the highest weights were too high, and for the lowest weights too low, it may be preferable to admit some imprecision among the studies with more nearly average weights in order to improve the precision of and reduce the discrepancy between the estimates of extremely high or low weights.

An advantage of both equations (10) and (11) for estimating the variance of the log rate ratio is that they rely only on the relationship between $\hat{Var}\ Y$ and $\hat{\mu}$, and not on their relationship to $\hat{p}$. Thus the necessary parameters for estimating weights for LnRRate for studies that provide only proportions can be estimated from a regression including all studies which provide means and variances, regardless of whether they also provide proportions. Nevertheless, these weights are estimated from quantities which estimate the variance that would have been obtained if the sample means and variances had been provided, rather than from *bona fide* estimates of variance.

The Delta Method

Finally, we also applied the delta method to estimate variances. Derivations are provided in Appendices 6 and 7.

Delta method to estimate weight for LnRRisk given means and variances

Figure 7 shows weights for LnRRisk (given means and variances) obtained by way of the delta method assuming Poisson, geometric, or negative binomial distributions (as contrasted against figure 5, which shows weights for LnRRisk by way of the substitution method). First we recall that substitution of the weight of the LnRRate yields estimates that are much smaller than the observed weight of the LnRRisk, for example by a factor of about 18 for the data point at the far right represented by a dash. By contrast,

the delta method assuming an underlying geometric distribution yields weights that are much greater, and the delta method assuming an underlying Poisson distribution yields weights that are vastly greater, than the observed weights. Again, taking the study at the farthest right as an example, the estimated weight, assuming a Poisson distribution, of $e^{28.7}$ is $e^{28.7-7.8} = e^{20.9}$ or over a billion times the observed weight. If a geometric distribution is assumed, then the estimated weight is $e^{12.1}$, which is $e^{12.1-7.8} = e^{4.3}$ or over 70 times the observed weight. The unrealistically large weights (and small variances) reflect the fact that these distributions are poorly suited to our count data, but if we assume a negative binomial distribution, then the estimated weight for this study (as well as all the other studies in this figure) is very close to the observed value for the LnRRate but not for the LnRRisk.

--- Figure 7 about here ---

Delta method to estimate weight for LnRRate given proportions

Figure 8 shows weights for LnRRate (given proportions) obtained by way of the delta method assuming Poisson or geometric distributions, or assuming the associations obtained by way of regression in Appendix 7 (as contrasted against figure 6, which shows weights for LnRRate by way of the substitution method). Recall that substitution of the weight of the LnRRisk yields estimates greater than the observed weight of the LnRRate, for example by a factor of 1346/149=9 for the data point at the far right represented by a dash. In this case, application of the delta method estimated weights closer to the observed values. In the study at the farthest right, the estimated weight, assuming a Poisson distribution, of $514 \left(\approx e^{6.2} \right)$ is about 3.5 times the observed weight of 149. If a geometric distribution is assumed, then the estimated weight is $224 \left(\approx e^{5.4} \right)$, which is only

51% greater than the observed weight. Weights estimated by the delta method assuming the associations obtained by way of regression were similar to those obtained under the assumption of a geometric distribution, for example $214\left(\approx e^{5.4}\right)$ for the study at the far right. Since we are given only one parameter, we did not attempt the negative binomial assumption, which would require two parameters.

--- Figure 8 about here ---

Summary of results from delta method

Estimates of weights from the delta method for LnRRisk assuming an underlying negative binomial distribution were so small that the contribution of such studies to a meta-analysis would be negligible. Estimates of weights for LnRRate assuming either a Poisson or a geometric distribution were greater than the observed values; therefore they would assign greater weight to studies that provided less information. Since neither of these results was satisfactory for the purpose of meta-analysis, we did not compare these weights to those obtained by other methods.

DISCUSSION

These derived methods for estimating risk ratios given only means and variances, or rate ratios given only proportions, are useful for meta-analysis of randomized trials, when risk ratios and rate ratios may be preferred over odds ratios and standardized mean differences.

Compared to simply interchanging LnRRisk with LnRRate, regression and the method of moments assuming a negative binomial distribution provided substantial improvements in estimates of LnRRisk. MSE was reduced from 0.34 to values of 0.05 or 0.09. Bias in estimates of LnRRisk estimates was small. By contrast, the same methods

yielded no improvement in precision of estimates of LnRRate over simple substitution of LnRRisk.

Estimated values of weight were greatly improved for both directions compared to naïve substitution between the LnRRisk weight and the LnRRate weight. MSE for simple substitution in either direction was 2.58 on the log scale in a sample with the median log weight LnRRisk at 3.80 and the median log weight LnRRate at 2.12. MSE for log weight LnRRisk was reduced to 0.10 by regression estimating proportion as a function of mean and variance and to 0.32 by the method of moments assuming a negative binomial distribution. MSE for log weight LnRRate was reduced to 0.32 by regression estimating both the mean and variance as functions of the proportion and to 0.38 by the method of moments assuming a negative binomial distribution.

Given these results, the best fit in either direction was obtained by way of regressions or by the method of moments assuming a negative binomial distribution. If a sufficient number of studies are available and the associations among means, variances, and proportions are strong, then regression may be the simpler solution to apply. The negative binomial fit can be quite complicated, particularly in the case of transformation from means and variances to proportions. Numeric solutions involving the regression assumption that Var[Y] is a function of E[Y] are necessary not only to estimate the VarLnRRate but even to estimate the LnRRate itself.

For estimation of LnRRisk and associated weights given only means, Var[Y], and sample sizes, the regression approach and the method of moments assuming a negative binomial distribution both provided minimal MSE and variance. The negative binomial approach can provide satisfactory estimates even if the research base does not provide

enough primary studies for confident estimation of regression parameters needed for equation (5).

For estimation of LnRRate and associated weights given only proportions and sample sizes, a simple ad hoc solution is to substitute LnRRisk for the missing LnRRate and to take advantage of the relationship shown in equation 10 or 11 for the weights. This strategy is particularly convenient because it does not require regression of E[Y] and Var[Y] on $\hat{p}$. Instead, only regression of LnVar[Y] on LnE[Y] is required. If the coefficient $\hat{\beta}_1$ for LnE[Y] is close to 2, then $\mu^{\hat{\beta}_1}/\mu^2$ can be evaluated at the geometric mean of $\hat{\mu}$ for estimating the variance by equation (11).

In the studies for which the individual-level data were available, there was a strong linear relationship between the log of the number of episodes of or partners for unprotected sex (plus one) and the log of the number of men giving that response. The log-linearity is not captured by any of the parametric distributions we have used. In fact, it suggests that the count data might be modeled well by the so-called zeta or zipf distribution. Zeta is characteristic of count distributions where discrete items are listed in descending order of frequency. For example, in typical English text, the most frequently occurring word is "the," followed by "a," "and," and "of" (reference). When listed in descending order, these frequencies tend to follow a zeta distribution.

However, there are several difficulties in trying to implement this approach. First, zeta distributions often do not admit a finite population mean or variance depending on the value of the underlying parameter. Any attempt to model the count data by a parametric distribution partly comprised of a zeta distribution may produce a parameter estimate which precludes a finite variance or a finite mean, rendering useless our

previous technique of expressing proportion as a function of population mean and variance. Second, the sample geometric mean, not the arithmetic mean, is the sufficient statistic for the zeta distribution parameter. Thus, studies reporting the arithmetic mean do not provide a summary statistic of any utility for a zeta count data model.

There are several limitations to the illustrations we have undertaken here. Perhaps most importantly, the estimates of variance which result from the substitution process are not actually the variance of the LnRRisk given only means and standard deviations, or of the LnRRate given only proportions. Instead they represent an attempt to estimate the variances that would have been observed for the LnRRisk if the proportion had been reported, or for the LnRRate if the means and standard deviations had been reported. In addition, some of the example studies used in these analyses were group-randomized trials; for this illustration we ignore the effect of intraclass correlation. Finally, these analyses focus on transformation only at a single time point; therefore they ignore baseline to follow-up correlations within each individual. Several approaches have been suggested for addressing correlation within subject across time and baseline differences in the outcome variable between treatment groups ([34] [35] [36]).

While the substitution approach for estimating variances of intervention effects has limitations, it did provide estimates which tended to closely approximate the observed values. A statistically valid approach to estimating variances should be to assume a parametric distribution for the count data and to apply the delta method by using the relationship between the mean and proportion specified by that distribution, but the delta method yielded variances (and weights) which did not approximate the observed values.

Consequently, we tried a variety of methods to estimate proportions from means and variances and vice versa. None was completely satisfactory: methods which appear to yield comparable weights for both sets of measures are not statistically valid, whereas attempts at a statistically rigorous approach gave weights for one metric which were not of the same numerical order as those for the other, making it difficult to combine the results in a meta-analysis.

A further complication occurred when we had to assume an underlying parametric distribution of our count data in order to estimate proportions from directly observed sample means and variances (and vice versa). Ignoring the fact that none of the workable distributions we explored accurately represents the count data in our studies, there is an inherent problem in that we are trying to compare a set of studies whose estimates and standard errors have been indirectly calculated using the delta method on the premise that the count data follow a specific parametric distribution, with another set of studies, whose metrics have been calculated by non-parametric techniques. This sometimes leads to the counter-intuitive result that the indirectly observed estimates are more efficient (have greater weight) than the directly observed ones. For example, in estimating proportions from means and variances, if we assume an underlying geometric distribution, the proportion estimates are necessarily functions of the sample mean, and the latter is the maximum likelihood estimator (MLE) for the population mean. It follows that the proportion estimate based on the sample mean is the MLE for the population proportion. As MLEs are asymptotically efficient, this estimator should outperform any other, in particular the non-parametric one (proportion of observed counts above the cut-off threshold), in terms of variance. Hence, the weights from this approach will be greater

than those for the substitution approach. Conversely, if we assume the count data follow

the two-parameter negative binomial distribution, then, unlike the sample mean, the

sample variance is not the MLE for its population counterpart. Accordingly, estimates of

proportion based on these are not efficient, and this is reflected in the fact that the

weights of these estimators are less than those of the non-parametric estimators.

Unfortunately, while the negative binomial has the appeal of producing good point

estimates and weights for the directly observed metrics, thus appearing to penalize the

indirectly estimated metrics with lower weights, the latter are so low as to have negligible

effect in a meta-analysis.

Delta methods have previously been used in meta-analysis for comparing the odds

ratio and standardized mean difference, based on the assumption that the data are normal

([37], [38]). A simple approximation to these methods is described in Chinn ([2]). Although

Chinn's method is often used for comparing odds ratios and standardized mean

differences in count data studies, we maintain that this practice is inappropriate because it

is an approximation to a delta method based on normality and because count data such as

ours are patently non-normal owing to their skewness. We could not perceive an analogy

to Chinn's approximation for our purposes of comparing risk ratios and rate ratios and we

resisted applying a delta method, as in Whitehead and Suissa, predicated on normality.

Chinn's method works because of the fortuitous coincidence that the standard logistic

distribution, which arises naturally in the context of odds ratio, is very close to a scaled

version of the standard normal. We could discern no comparable link in the case where

LnRRisk is the measure of interest. In one sense, the direct analogy to the work of

Whitehead and Suissa is our negative binomial delta method. We note that in simulations

of normal data (not shown here), we found the Whitehead and Chinn weights for odds

ratios calculated from observed means and variances considerably more efficient than the

non-parametric weights for directly observed odds ratios, depending on the cut-point.

This is not apparent in Chinn's paper as cutpoints are ignored for simplicity, but it does

reflect our previous observation that estimates based on MLEs will be more efficient than

alternatives.

The substitution approach has the appeal of yielding comparable weights in both

sets of studies. Although it is arguably in the spirit of meta-analysis, it has no statistical

justification. It calculates point estimates by assuming an underlying parametric

distribution, yet estimates standard errors by applying the non-parametric method. Thus,

it is neither coherent in its association of standard error with estimate, nor does it penalize

those indirectly calculated estimates with larger standard errors. We leave it to the reader

to judge its merits.

The regression techniques suffer the same inconsistency and lack of statistical

rigor. However, they represent an empirical attempt to reproduce the count data patterns

from those studies providing both proportions and means and variances.

Results for the Poisson distribution under the method of moments correspond to

equations previously shown ([39]). Our results add the geometric and negative binomial

distributions to the list of transformation formulas. In other analyses not presented here,

we also derived formulas by way of the method of moments assuming the zero-inflated

Poisson and zero-inflated geometric distributions, but results were less satisfactory. In

zero-inflated models, some proportion of subjects are assumed not to be at risk, while the

rest of the subjects take on outcome values following a count distribution such as

Poisson, geometric, or negative binomial [40 41]. Among the three 2-parameter distributions considered (negative binomial [*not* zero-inflated], zero-inflated Poisson, and zero-inflated geometric), the negative binomial distribution was the most satisfactory, presumably because it is the most flexible. Zero-inflated models require that the direction of discrepancy from the Poisson or geometric models be toward excess zeros; instead the greater degree of discrepancy from the negative binomial model in our data was toward a small number of very large values. We did not consider the zero-inflated negative binomial distribution because it requires three input parameters, and the available data for these scenarios include only two (the mean and variance). The third moment, the skewness, is rarely if ever reported; therefore solutions based on the zero-inflated negative binomial distribution would not be useful.

This brings us to a practical observation that HIV prevention interventions for MSM might do well to focus more effort toward the small proportion of men who report the most unprotected sex. We also note that for any given number of occasions of unprotected sex, the capacity of a community to maintain an epidemic would seem to be lower if most members are mutually monogamous than if each unprotected occasion is with a different partner. For this reason, the number of different partners for any unprotected sex has substantial intuitive appeal as a primary outcome target for these interventions.

In conclusion, it would be most beneficial for researchers to measure and report dichotomous unprotected sex as well as number of occasions of and number of partners for unprotected sex as three separate outcomes. Similar precautions likely apply to other areas of health promotion that focus on reducing a number of risk events toward zero,

such as smoking cessation (number of cigarettes) and drug abuse treatment (number of injections). When the wider range of preferred information is not provided, some method of substitution of the available but limited information may be necessary if information from all studies is to be combined and compared in a single meta-analysis, in which case the methods described above may be useful.

Table 1. Four types of outcome metrics, either for randomized controlled trials or as required by the Chinn method.

	Type of Outcome	
	Dichotomies	Means
RCT	$Risk\ Ratio = \dfrac{p_1}{p_2}$	$Rate\ Ratio = \dfrac{\mu_1}{\mu_2}$
Chinn method	$Odds\ Ratio = \dfrac{p_1(1-p_2)}{p_2(1-p_1)}$	$SMD = \dfrac{\mu_1 - \mu_2}{\sigma}$

RCT = randomized controlled trial

SMD = standardized mean difference

p_1 = proportion reporting 1 or more occasions in the intervention condition

p_2 = proportion reporting 1 or more occasions in the comparison condition

μ_1 = mean number of occasions in the intervention condition

μ_2 = mean number of occasions in the comparison condition

σ = pooled standard deviation of occasions in both conditions

Table 2. Mean, variance, probability function (pf), and estimate $\hat{p}$ of the proportion at high risk given only observed mean $\overline{Y}$ and variance S^2 under three distributional assumptions. Support for all distributions is y = 0, 1, 2 …. Expressions for mean, variance, and pmf are given; derivation of estimates of $\hat{p}$ is provided in Appendix 2.

Distribution	Mean	Variance	Probability function	Proportion $\hat{p}$ at high risk	Where
Poisson	μ	μ	$\dfrac{\mu^y e^{-\mu}}{y!}$	$1 - e^{-\hat{\mu}}$	$\hat{\mu} = \overline{Y}$
Geometric	$\dfrac{1-g}{g}$	$\dfrac{1-g}{g^2}$	$g(1-g)^y$	$1 - \hat{g}$	$\hat{g} = \dfrac{1}{1+\overline{Y}}$
Negative binomial	$\dfrac{k(1-g)}{g}$	$\dfrac{k(1-g)}{g^2}$	$\dfrac{(y+k-1)!}{y!(k-1)!} g^k (1-g)^y$	$1 - \hat{g}^{\hat{k}}$	$\hat{g} = \dfrac{\overline{Y}}{S^2}$ $\hat{k} = \dfrac{\overline{Y}^2}{S^2 - \overline{Y}}$

Table 3: Estimation of mean and variance given only observed proportion $\hat{p}$ at high risk.

Derivations are provided in Appendix 4.

Distribution	Mean	Variance
Poisson	$-\ln(1-\hat{p})$	$-\ln(1-\hat{p})$
Geometric	$\dfrac{\hat{p}}{1-\hat{p}}$	$\dfrac{\hat{p}}{(1-\hat{p})^2}$
Negative binomial*	$\hat{\mu}$ where $\ln(1-\hat{p}) = \dfrac{\hat{\mu}^2\left[(1-b)\ln\hat{\mu} - a\right]}{e^a\hat{\mu}^b - \hat{\mu}}$	$e^a\hat{\mu}^b$

* with the further assumption that $LnVar[Y] = bLnE[Y] + a$ in order to reduce the 2-parameter negative binomial distribution to 1 parameter.

Table 4. Mean squared error (MSE) and bias in estimates of LnRRisk and log of weight

of LnRRisk given means and variances but not given proportions at 79 time points in 19

studies of HIV prevention for men who have sex with men.

Method for estimating proportion	LnRRisk		Log of weight LnRRisk	
	MSE	Bias	MSE	bias
Identity*	0.34	0.02	2.58	-1.47
Regression**	0.05	0.03	0.10	-0.001
Method of moments				
Poisson	0.13	0.10	32.81	4.61
Geometric	0.11	0.09	3.17	1.64
Negative Binomial	0.09	0.01	0.15	-0.09

* Assuming that LnRRisk = LnRRate and VarLnRRisk = VarLnRRate

** Assuming that $\ln E[Y] = 1.13 \times \text{logit}(p) + 1.76$.

Table 5. Mean squared error (MSE) and bias in estimates of LnRRate and log of weight

of LnRRate at 79 time points in 19 studies of HIV prevention for men who have sex with

men.

Method for estimating LnRRate from $\hat{p}_1$ and $\hat{p}_2$	LnRRate		Log of weight of LnRRate	
	MSE	bias	MSE	Bias
Identity*	0.34	-0.02	2.58	1.47
Regression**	0.34	-0.08	0.32	0.06
Method of moments				
Poisson	0.33	-0.04	1.82	1.22
Geometric	0.34	-0.07	1.28	0.99
Negative Binomial***	0.36	-0.02	0.38	0.07

* Assuming that LnRRate = LnRRisk and VarLnRRate = VarLnRRisk

** Assuming that $\ln E[Y] = 1.13 \times \text{logit}(p) + 1.76$ and $\ln Var[Y] = 1.80 \times \text{logit}(p) + 5.34$

*** Assuming that $Var[Y] = e^{2.17} E[Y]^{1.93}$

Figure Titles

Figure 1. Distribution of count outcomes at baseline in eight example studies of HIV prevention for men who have sex with men (up to count of 19) and sum across all eight studies (up to count of 150), with smoothed running averages in intervals.

Figure 2. Logit of proportion and natural log of variance by natural log of mean in 158 combinations of treatment condition by time in 19 studies of HIV prevention for men who have sex with men.

Figure 3. Observed LnRRate and estimates of LnRRisk by observed LnRRisk from 23 comparisons in 19 studies of HIV prevention for men who have sex with men (two additional data points are outside the scale of this figure).

Figure 4. Natural log of observed weight of LnRRate and estimates of weight of LnRRisk by observed weight of LnRRisk at 79 time points in 19 studies of HIV prevention for men who have sex with men.

Figure 5. Observed LnRRisk and estimates of LnRRate by observed LnRRate from 23 comparisons in 19 studies of HIV prevention for men who have sex with men (two additional data points are outside the scale of this figure).

Figure 6. Natural log of observed weight of LnRRisk and estimates of weight of LnRRate by observed weight of LnRRate at 79 time points in 19 studies of HIV prevention for men who have sex with men.

Figure 7. Natural log of observed weight of LnRRate and estimates of weight of LnRRisk (from the delta method) by observed weight of LnRRisk at 79 time points in 19 studies of HIV prevention for men who have sex with men.

Figure 8. Natural log of observed weight of LnRRisk and estimates of weight of LnRRate (from the delta method) by observed weight of LnRRate at 79 time points in 19 studies of HIV prevention for men who have sex with men.

Appendix 1: Estimating the usual variance of LnRRisk and LnRRate by the delta method

Let X be a random variable with sample size n for which

$$\sqrt{n}(X - \theta) \xrightarrow{d} N(0, \sigma^2)$$

where θ and σ^2 are finite valued constants and $\xrightarrow{d}$ denotes convergence in distribution as sample size n increases. By the delta method:

$$\sqrt{n}[g(X) - g(\theta)] \xrightarrow{d} N(0, \sigma^2 [g'(\theta)]^2).$$

The variance of $g(X)$ with sample size n can then be estimated as $\sigma^2 [g'(\theta)]^2 / n$.

Binomial distribution and the variance of LnRRisk given sample proportions

Suppose X is Binomial with parameters p and n. Since

$$\sqrt{n}\left(\frac{X}{n} - p\right) \xrightarrow{d} N(0, p(1 - p))$$

Then, by the delta method with $g(\theta) = \log(\theta)$, then $g'(\theta) = 1/\theta$ and

$$\sqrt{n}\left[\log\left(\frac{X}{n}\right) - \log(p)\right] \xrightarrow{d} N(0, p(1 - p)(1/p)^2) = N\left(0, \frac{1-p}{p}\right),$$

the variance of $\log\left(\frac{X}{n}\right)$ is approximately $\frac{1-p}{np}$.

If $\hat{p}_1$ and $\hat{p}_2$ are estimates from independent samples of sizes n_1 and n_2 respectively, then the logarithm of the estimated relative risk $\hat{p}_1/\hat{p}_2$ is approximately normally distributed with variance that can be estimated by

$$\frac{1-\hat{p}_1}{n_1\hat{p}_1}+\frac{1-\hat{p}_2}{n_2\hat{p}_2}.$$

The meta-analytical weight is then the reciprocal of this variance:

$$\frac{1}{\dfrac{1-\hat{p}_1}{n_1\hat{p}_1}+\dfrac{1-\hat{p}_2}{n_2\hat{p}_2}}$$

The variance of LnRRate given sample means and variances

Suppose X is a count variable of an unspecified distribution with mean μ and variance σ^2. By the Delta method with $g(\theta)=\log(\theta)$, then $g'(\theta)=1/\theta$ and

$$\sqrt{n}\left[\log\left(\frac{X}{n}\right)-\log(\mu)\right]\xrightarrow{d} N\left(0,\frac{\sigma^2}{\mu^2}\right),$$

the variance of $\log\left(\dfrac{X}{n}\right)$ is approximately $\dfrac{Var[X]}{n\mu^2}$.

If $\hat{\mu}_1$ and $\hat{\mu}_2$ are estimates from independent samples of sizes n_1 and n_2 respectively, then the logarithm of the estimated relative rate $\hat{\mu}_1/\hat{\mu}_2$ is approximately normally distributed with variance that can be estimated by

$$\frac{V\hat{a}r[X_1]}{n_1\hat{\mu}_1^2}+\frac{V\hat{a}r[X_2]}{n_2\hat{\mu}_2^2}.$$

The meta-analytical weight is then the reciprocal of this variance:

$$\frac{1}{\dfrac{V\hat{a}r[X_1]}{n_1\hat{\mu}_1^2} + \dfrac{V\hat{a}r[X_2]}{n_2\hat{\mu}_2^2}}$$

Appendix 2. **Expressing the parameters of each distribution in terms of mean and variance**

In the first series of scenarios, we are given the mean $\bar{Y}$ and variance

$S^2 = \dfrac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ and we need to estimate the proportion $\hat{p}$. Expressions for the mean

and variance in terms of the parameters of each candidate distribution are well known and

are provided in Table 2. The following steps show the algebra used to reverse these

formulas to express the parameters of each distribution in terms of the mean and

variance, as shown in the far right column of Table 2.

<u>Poisson distribution</u>:

    From $E[Y] = \mu$, we obtain $p = 1 - e^{-\mu} = 1 - e^{-E[Y]}$

<u>Geometric distribution</u>:

    From $E[Y] = \dfrac{1-g}{g}$, we obtain $g = \dfrac{1}{E[Y]+1}$ and substitute into $p = 1 - g$

<u>Negative binomial distribution</u>:

First express g in terms of $E[Y]$ and $Var[Y]$:

$$E[Y] = \frac{k(1-g)}{g} \hspace{5cm} \#A1$$

$$Var[Y] = \frac{k(1-g)}{g^2} \hspace{5cm} \#A2$$

#A1 divided by #A2 gives $g = \dfrac{E[Y]}{Var[Y]}$ $\hspace{4cm}$ #A3

Now express k in terms of $E[Y]$ and $Var[Y]$. From #A1 and #A3:

$$k = \frac{g}{(1-g)} \times E[Y] = \left(\frac{E[Y]}{Var[Y] - E[Y]}\right) E[Y] = \frac{E^2[Y]}{Var[Y] - E[Y]} \qquad \#A4$$

These expressions for g and k can then be substituted into the expression $p = 1 - g^k$

Appendix 3. **Expressing method of moments estimates of proportion in terms of mean and variance under each assumed distribution**

We define the population proportion $p = 1 - P[Y = 0]$. For each distribution, the population proportion p has now been expressed in terms of $E[Y]$ and $Var[Y]$ in Appendix 1. Now the method of moments estimator $\hat{p}$ for proportion is obtained by replacing $E[Y]$ and $Var[Y]$ in those expressions for p by their method of moments estimators, $\overline{Y}$ and S^2 respectively to obtain the expression for $\hat{p}$ for each distribution as shown in the rightmost column of Table 2.

Poisson distribution:

From appendix 1 we have $p = 1 - e^{-\mu}$ where $\mu = E[Y]$

$$\text{Therefore } \hat{p} = 1 - e^{-\hat{\mu}} \text{ where } \hat{\mu} = \overline{Y}$$

Geometric distribution:

From appendix 1 we have $p = 1 - g$ where $g = \dfrac{1}{E[Y] + 1}$

$$\text{Therefore } \hat{p} = 1 - \hat{g} \text{ where } \hat{g} = \dfrac{1}{\overline{Y} + 1}$$

Negative binomial distribution:

From appendix 1 we have $p = 1 - g^k$ where $g = \dfrac{E[Y]}{Var[Y]}$ and $k = \dfrac{E^2[Y]}{Var[Y] - E[Y]}$

$$\text{Therefore } \hat{p} = 1 - \hat{g}^{\hat{k}} \text{ where } \hat{g} = \dfrac{\overline{Y}}{S^2} \text{ and } \hat{k} = \dfrac{\overline{Y}^2}{S^2 - \overline{Y}}$$

Appendix 4. Expressing the population mean and variance in terms of proportion p

In the second series of scenarios, we are given the proportion $\hat{p}$ and we need to estimate

the mean and variance. We begin with the expressions for the population proportion p

and the parameters of each distribution in terms of the mean and variance as shown in the

two rightmost columns of Table 2. The following steps show the algebra used to reverse

these formulas to express mean and variance in terms of proportion.

Poisson distribution: Given $p = 1 - e^{-\mu}$ where $\mu = E[Y]$,

      then $\ln(1 - p) = -\mu$ and therefore

      $E[Y] = Var[Y] = \mu = -\ln(1 - p)$

Geometric distribution: Given $p = 1 - g$ where $g = \dfrac{1}{E[Y]+1}$, then

      $E[Y] = \dfrac{1-g}{g} = \dfrac{p}{1-p}$ and

      $Var[Y] = \dfrac{1-g}{g^2} = \dfrac{p}{(1-p)^2}$

Negative binomial distribution:

From $p = 1 - g^k$ where $g = \dfrac{E[Y]}{Var[Y]}$ and $k = \dfrac{E^2[Y]}{Var[Y] - E[Y]}$,

we obtain $1 - p = g^k$, so that

$$Ln(1 - p) = kLn(g) = \frac{E^2[Y]}{Var[Y] - E[Y]}Ln\left(\frac{E[Y]}{Var[Y]}\right) = \frac{E^2[Y][Ln(E[Y]) - Ln(Var[Y])]}{Var[Y] - E[Y]}$$

If we assume $Var[Y] = e^a E^b[Y]$ with a and b assumed known, or estimated from

regression parameters as described in the text, then:

$$Ln(1-p) = \frac{E^2[Y][(1-b)Ln(E[Y])-a]}{e^a E^b[Y] - E[Y]}$$

There is no closed algebraic form for E[Y] from this last formula, so solutions must be

obtained numerically.

Appendix 5. Expressing the population mean and variance in terms of the population proportion p, and then present the method of moments estimators of these by replacing p by $\hat{p}$

Poisson distribution:

From appendix 3 we have $E[Y] = Var[Y] = \mu = -\ln(1-p)$

Therefore

$$\hat{E}[Y] = V\hat{a}r[Y] = \hat{\mu} = -\ln(1-\hat{p})$$

Geometric distribution:

From appendix 3 we have

$$E[Y] = \frac{1-g}{g} = \frac{p}{1-p}$$

$$Var[Y] = \frac{1-g}{g^2} = \frac{p}{(1-p)^2}$$

Therefore

$$\hat{E}[Y] = \frac{1-\hat{g}}{\hat{g}} = \frac{\hat{p}}{1-\hat{p}}$$

$$V\hat{a}r[Y] = \frac{1-\hat{g}}{\hat{g}^2} = \frac{\hat{p}}{(1-\hat{p})^2}$$

Negative binomial distribution:

From appendix 3 we have

$$Ln(1-p) = \frac{E^2[Y][(1-b)Ln(E[Y])-a]}{e^a E^b[Y] - E[Y]}$$

Thefore

$$Ln(1-\hat{p}) = \frac{\overline{Y}^2\left[(1-b)Ln(\overline{Y})-a\right]}{e^a\overline{Y}^b - \overline{Y}}$$

A numeric solution must then be obtained for $\overline{Y}$ given $\hat{p}$

Appendix 6: Delta method to estimate variance of LnRRisk given means and variances

Here we apply the delta method to estimate the variance of the LnRRisk as described in Appendix 1, but with the assumption that we are given only means, variances (instead of proportions), and sample sizes. The underlying count variable is assumed to be distributed as either Poisson, geometric, or negative binomial, and the proportion at higher risk $\hat{p}$ is estimated under each distribution as shown in Table 2.

Estimation of the variance of LnRRisk assuming Poisson distribution of the underlying count variable

We want to estimate the variance of $\ln p = \ln(1 - P[Y = 0]) = \ln(1 - e^{-\lambda}) = f(\lambda)$

By invariance, the MLE is $\ln(1 - e^{-\hat{\lambda}})$ which is estimated as $\ln(1 - e^{-\bar{Y}})$

For the delta method we take the function $f(t) = \ln(1 - e^{-t})$

The derivative of this function is $f'(\lambda) = \dfrac{1}{1 - e^{-\lambda}} \times e^{-\lambda} = \dfrac{1}{e^{\lambda} - 1}$

So by the delta method:

$$\sqrt{n}\left(f(\hat{\lambda}) - f(\lambda)\right) \xrightarrow{d} N\left(0, \lambda[f'(\lambda)]^2\right)$$

$$\sqrt{n}\left[\ln(1 - e^{\hat{\lambda}}) - \ln(1 - e^{\lambda})\right] \xrightarrow{d} N\left(0, \frac{\lambda}{(e^{\lambda} - 1)^2}\right)$$

Thus the variance of $\ln\left(1 - e^{-\hat{\lambda}}\right)$ is approximately $\dfrac{\lambda}{n\left(e^{\lambda} - 1\right)^2}$

which is estimated by $\dfrac{\overline{Y}}{n\left(e^{\overline{Y}} - 1\right)^2}$

Estimation of the variance of LnRRisk assuming geometric distribution of the underlying count variable

We want to estimate the variance of

$$\ln p = \ln\left(1 - P[Y = 0]\right) = \ln\left(1 - g\right)$$

By invariance, the MLE is $\ln\left(1 - \hat{g}\right)$ which is estimated as $\ln\left(\dfrac{\overline{Y}}{\overline{Y} + 1}\right)$ (from Table 2)

We define proportion $= \pi = 1 - g = 1 - \dfrac{1}{\mu + 1} = \dfrac{\mu}{\mu + 1}$

$$\ln \pi = \ln \mu - \ln\left(\mu + 1\right)$$

The variance of the geometric distribution is $\sigma^2 = \dfrac{1 - g}{g^2} = \dfrac{\pi}{\left(1 - \pi\right)^2} = \mu\left(\mu + 1\right)$

For the delta method we take the function

$$f(t) = \ln\left(\dfrac{t}{t+1}\right) = \ln(t) - \ln(t + 1)$$

$$f'(t) = \dfrac{1}{t} - \dfrac{1}{t+1} = \dfrac{1}{t\left(t+1\right)}$$

So by the delta method:

$$\sqrt{n}\left(\overline{Y} - \mu\right) \xrightarrow{d} N\left(0, \sigma^2\right) \qquad\qquad \sigma^2 = \mu\left(\mu + 1\right)$$

$$\hat{\pi} = f\left(\overline{Y}\right) = \ln\frac{\overline{Y}}{\overline{Y} + 1}$$

$$\sqrt{n}\left(\ln\hat{\pi} - \ln\pi\right) \xrightarrow{d} N\left(0, \left[f'(\mu)\right]^2 \sigma^2\right)$$

Thus the variance of $\ln\hat{\pi}$ is approximately $= \dfrac{1}{n} \times \dfrac{1}{\left[\mu\left(\mu + 1\right)\right]^2} \mu\left(\mu + 1\right) = \dfrac{1}{n\mu\left(\mu + 1\right)}$

which is estimated by $\dfrac{1}{n\overline{Y}\left(\overline{Y} + 1\right)}$

Estimation of the variance of LnRRisk assuming negative binomial (NB) distribution of the underlying count variable

From standard theory ([42]),

$$\sqrt{n}\begin{bmatrix} \overline{Y} - \mu \\ S^2 - \sigma^2 \end{bmatrix} \xrightarrow{d} N_2\left(0, \Sigma\right),$$

$$\text{where } \Sigma = \begin{bmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_3 - \sigma^4 \end{bmatrix}$$

and $\mu_k = E\left[(Y - \mu)^k\right]$ is the kth central moment. As the result is asymptotic, S^2 can be either the unbiased sample variance or the $1/n$ version.

For NB with support $\{0, 1, 2, \ldots\}$ and mass function

$$P[Y = y] = \frac{\Gamma(\theta + y)}{\Gamma(\theta)y!} \frac{\theta^\theta \mu^y}{(\theta + \mu)^{\theta + y}},$$

with $p = \theta/(\theta + \mu)$, we have $E[Y] = \mu = \theta q/p$, $Var[Y] = \sigma^2 = \mu + \mu^2/\theta = \theta q/p^2$.

The skewness is

$$\frac{E\left[(Y-\mu)^3\right]}{\sigma^3} = \frac{2-p}{\sqrt{\theta q}},$$

giving

$$\mu_3 = \mu\left(1+\frac{\mu}{\theta}\right)\left(1+\frac{2\mu}{\theta}\right).$$

The excess kurtosis is

$$\frac{E\left[(Y-\mu)^4\right]}{\sigma^4} - 3 = \frac{6}{\theta} + \frac{p^2}{\sqrt{\theta q}},$$

giving

$$\mu_4 - \sigma^4 = \mu\left(1+\frac{\mu}{\theta}\right)\left[1+2\mu\left(1+\frac{1}{\theta}\right)\left(1+\frac{3}{\theta}\right)\right].$$

It follows that, for NB, the above asymptotic result takes the form

$$\sqrt{n}\left[\begin{array}{c}\overline{Y}-\mu\\S^2-\sigma^2\end{array}\right] \xrightarrow{d} N_2(0,\Sigma),$$

where $\Sigma = \mu\left(1+\dfrac{\mu}{\theta}\right)\left[\begin{array}{cc} 1 & 1+\dfrac{2\mu}{\theta} \\[2mm] 1+\dfrac{2\mu}{\theta} & 1+2\mu\left(1+\dfrac{1}{\theta}\right)\left(1+\dfrac{3}{\theta}\right)\end{array}\right]$

$$= \sigma^2\left[\begin{array}{cc} 1 & \dfrac{2\sigma^2}{\mu}-1 \\[3mm] \dfrac{2\sigma^2}{\mu}-1 & 1+2\sigma^2\left[1+\dfrac{3\left(\sigma^2-\mu\right)}{\mu^2}\right]\end{array}\right]$$

Now $P[Y = 0] = \left(\dfrac{\theta}{\theta + \mu}\right)^{\theta}$ and the method of moments estimates for the

parameters are

$$\hat{\mu} = \overline{Y}, \quad \hat{\theta} = \dfrac{\overline{Y}^2}{S^2 - \overline{Y}}.$$

Re-expressed, the proportion is

$$P[Y > 0] = 1 - \left(\dfrac{\mu}{\sigma^2}\right)^{\mu^2/(\sigma^2 - \mu)},$$

so that the MM-derived estimator for proportion $\pi = P[Y > 0]$ is

$$\hat{\pi} = 1 - \left(\dfrac{\overline{Y}}{S^2}\right)^{\overline{Y}^2/(S^2 - \overline{Y})}.$$

To get the asymptotic variances of the log of this estimator, we apply the delta method

with the function

$$f(u,v) = \log\left(1 - \left(\dfrac{u}{v}\right)^{u^2/(v-u)}\right)$$

to get

$$\sqrt{n}\left(\log \hat{\pi} - \log \pi\right) \xrightarrow{d} N_1\left(0, \tau^2\right),$$

$$\tau^2 = [f_u, f_v] \Sigma \begin{bmatrix} f_u \\ f_v \end{bmatrix},$$

where $f_u = \dfrac{\partial f}{\partial u}$ etc, partial derivatives are evaluated at $(u,v) = (\mu, \sigma^2)$. Thus, τ^2/n is the

asymptotic variance of $\log \hat{\pi}$.

The partial derivatives are:

$$f_u = \frac{1}{1-\left(\dfrac{u}{v}\right)^{u^2/(v-u)}} \times \frac{\partial}{\partial u}\left\{-\left(\frac{u}{v}\right)^{u^2/(v-u)}\right\}$$

$$= \frac{-\left(\dfrac{u}{v}\right)^{u^2/(v-u)}}{1-\left(\dfrac{u}{v}\right)^{u^2/(v-u)}} \times \frac{\partial}{\partial u}\left\{\frac{u^2}{v-u}\left(\log u - \log v\right)\right\}$$

$$= \left(1 - \frac{1}{1-\left(\dfrac{u}{v}\right)^{u^2/(v-u)}}\right)\left(\frac{u(2v-u)}{(v-u)^2}\left(\log u - \log v\right) + \frac{u}{v-u}\right)$$

$$f_v = \frac{-\left(\dfrac{u}{v}\right)^{u^2/(v-u)}}{1-\left(\dfrac{u}{v}\right)^{u^2/(v-u)}} \times \frac{\partial}{\partial v}\left\{\frac{u^2}{v-u}\left(\log u - \log v\right)\right\}$$

$$= \left(1 - \frac{1}{1-\left(\dfrac{u}{v}\right)^{u^2/(v-u)}}\right)\left(-\frac{u^2}{(v-u)^2}\left(\log u - \log v\right) - \frac{u^2}{v(v-u)}\right)$$

Evaluated at $(u,v) = (\mu, \sigma^2)$, these become

$$f_u = (1 - \pi^{-1})\left(\frac{\mu(2\sigma^2 - \mu)}{(\sigma^2 - \mu)^2}\left(\log\frac{\mu}{\sigma^2}\right) + \frac{\mu}{\sigma^2 - \mu}\right)$$

$$f_v = (1 - \pi^{-1})\left(-\frac{\mu^2}{(\sigma^2 - \mu)^2}\left(\log\frac{\mu}{\sigma^2}\right) - \frac{\mu^2}{\sigma^2(\sigma^2 - \mu)}\right)$$

In summary, the asymptotic variance τ^2/n of $\log\hat{\pi}$ is given by

$$\tau^2 = [f_u, f_v] \Sigma \begin{bmatrix} f_u \\ f_v \end{bmatrix} = f_u^2 A + 2 f_u f_v B + f_v^2 C,$$

where

$$A = \sigma^2$$

$$B = \sigma^2 \left(\frac{2\sigma^2}{\mu} - 1 \right)$$

$$C = \sigma^2 \left(1 + 2\sigma^2 \left[1 + \frac{3(\sigma^2 - \mu)}{\mu^2} \right] \right)$$

and f_u, f_v are as above.

To estimate τ^2 from the data, calculate $\hat{\tau}^2$ by replacing (μ, σ^2) by the method of moments estimator $(\bar{Y}, S^2)$.

We did not investigate the delta method for estimation of proportions from the means and variances from regression, because it would require distributional assumptions on the sample mean and variance and would therefore involve another two-dimensional delta method.

Empirical application of these three results (Poisson, geometric, and negative binomial) is shown in Figure 7.

Appendix 7: Delta method to estimate variance of LnRRate given proportions

Here we apply the delta method as described in Appendix A, but with the assumption that we are given only proportions and sample sizes for each condition. For the first derivation below, we assume a linear association between $\ln \mu$ and $\text{logit}(p)$. For the second and third derivations, the underlying count variable is assumed to be distributed as either Poisson or geometric, and the mean and variance are estimated from the proportion under each distribution as shown in Table 3. We did not conduct this process with an assumption of an underlying negative binomial distribution because that process requires two input parameters and we have only one (the proportion).

Estimation of the variance of LnRRate assuming a linear association between $\ln \mu$ and logit (p)

If $\ln \mu = a \times \text{logit}(p) + b$

(for example equation 6 indicates that a=1.13 and b=1.76 in the studies we located), and assuming this relationship is the same for both control and intervention groups, then

$$\ln RR = \ln \mu_I - \ln \mu_C = a \times \left[\text{logit}(p_I) - \text{logit}(p_C)\right] = a \times \ln OR$$

(the *b*s cancel out)

So an interesting consequence of this linear assumption is that LnRRate is a constant factor *a* times the log odds ratio (LnOR). Since *a* is constant, the delta method asymptotic variance estimate is therefore

$$VarLnRR = a^2 \times VarLnOR$$

$$= a^2 \left(\frac{1}{n_I p_I} + \frac{1}{n_I(1 - p_I)} + \frac{1}{n_C p_C} + \frac{1}{n_C(1 - p_C)} \right)$$

in other words, a^2 times the usual sum of the reciprocals of the cell counts.

Estimation of the variance of LnRRate assuming Poisson distribution of the underlying count variable

We need a variance for $\ln \mu$

By the Central Limit Theorem (CLT):

$$\sqrt{n}(\hat{p} - \pi) \overset{d}{\to} d(0, \pi(1 - \pi)) \quad \text{where} \quad \hat{p} = \sum_i I[Y_i > 0]/n$$

Let $f(t) = \ln[-\ln(1 - t)]$ so that $f'(t) = \dfrac{1}{-\ln(1 - t)} \times \dfrac{1}{1 - t}$

and $f(\pi) = \ln \mu$ since $\mu = -\ln(1 - \pi)$.

By delta method: $\sqrt{n}(f(\hat{p}) - f(\pi)) \overset{d}{\to} N\left(0, [f'(\pi)]^2 \pi(1 - \pi)\right)$

So $\sqrt{n}(\ln(-\ln(1 - p)) - \ln \mu) \overset{d}{\to} N\left(0, \left(\dfrac{1}{-\ln(1 - \pi)} \times \dfrac{1}{1 - \pi}\right)^2 \times \pi(1 - \pi)\right)$

Thus the variance of $\ln(\mu)$ is approximately

$$\left(\dfrac{1}{-\ln(1 - \pi)}\right)^2 \times \dfrac{\pi}{1 - \pi}$$

which can be estimated by $\left(\dfrac{1}{-\ln(1 - \hat{p})}\right)^2 \times \dfrac{\hat{p}}{1 - \hat{p}}$

Estimation of the variance of LnRRate assuming geometric distribution of the underlying count variable

Define proportion $= \pi = 1 - g$

$\pi = \dfrac{\mu}{\mu + 1}$ so $\mu = \dfrac{\pi}{1 - \pi}$

$$\ln(\mu) = \ln(\pi) - \ln(1 - \pi)$$

$$f(t) = \ln(t) - \ln(1-t) \qquad f'(t) = \frac{1}{t} + \frac{1}{1-t} = \frac{1}{t(1-t)}$$

By CLT:

$$\sqrt{n}(\hat{p} - \pi) \xrightarrow{d} N(0, \sigma^2) \qquad\qquad \sigma^2 = \pi(1-\pi)$$

$$\hat{\mu} = f(\hat{p}) = \ln \frac{\hat{p}}{1-\hat{p}}$$

$$\sqrt{n}(\ln \hat{\pi} - \ln \pi) \to d \to N\left(0, [f'(\mu)]^2 \sigma^2\right)$$

Thus the asymptotic variance is approximately

$$\frac{1}{n} \times [f'(\pi)]^2 \times \sigma^2 = \frac{1}{n} \times \left[\frac{1}{\pi(1-\pi)}\right]^2 \times \pi(1-\pi) = \frac{1}{n\pi(1-\pi)}$$

which can be estimated by $\dfrac{1}{n\hat{p}(1-\hat{p})}$

Application of these three results (regression, Poisson, and geometric) is shown in Figure 8.

REFERENCES

[1] Hasselblad V, Hedges LV. Meta-analysis of screening and diagnostic tests. *Psych Bulletin* 1995;**117**:167-178.

[2] Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine* 2000;**19**:3127-3131.

[3] Johnson WD, Semaan S, Hedges LV, Ramirez G, Mullen PD, Sogolow E. A protocol for the analytical aspects of a systematic review of HIV prevention research. *J Acquir Immune Defic Syndr* 2002;**30**:S62-S72.

[4] Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 (updated September 2008). The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

[5] Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statist Med* 2002;**21**:1575–1600.

[6] Altman DG, Deeks JJ. Odds ratios should be avoided when events are common. *BMJ* 1998;**317**:1318.

[7] Katz KA. The (relative) risks of using odds ratios. *Arch Dermatol* 2006;**142**:761-764.

[8] Deeks JJ. Odds ratios should be used only in case-control studies and logistic regression analyses. *BMJ* 1998;**317**:1155-1156.

[9] Bracken MB, Sinclair JC. Avoidable systematic error in estimating treatment effects must not be tolerated. *BMJ* 1998;**317**:1156.

[10] Wackerly DD, Mendenhall W, Scheaffer RL. *Mathematical Statistics with Applications*. 5th edn. Duxbury Press, Wadsworth Publishing Company: Belmont, California, 1996.

[11] Oehlert GW. A note on the delta method. *The American Statistician* 1992;**46**:27-29.

[12] Johnson WD, Holtgrave DR, McClellan WM, Flanders WD, Hill AN, Goodman M. HIV intervention research for men who have sex with men: a 7-year update. *AIDS Education and Prevention* 2005;**17(6)**:568–589.

[13] Carballo-Diéguez A, Dolezal C, Leu C-S, Nieves L, Díaz F, Decena C, Balan I. A randomized controlled trial to test an HIV-prevention intervention for Latino gay and bisexual men: Lessons learned. *AIDS Care* 2005;**17**:314–328.

[14] EXPLORE Study Team. Effects of a behavioural intervention to reduce acquisition of HIV infection among men who have sex with men: the EXPLORE randomised controlled study. *Lancet* 2004;**364**:41–50.

[15] Flowers P, Hart GJ, Williamson LM, Frankis JS, Der GJ. Does bar-based, peer-led sexual health promotion have a community-level effect amongst gay men in Scotland? *Intl J STD & AIDS* 2002;**13**:102–108.

[16] Gold RS, Rosenthal DA. Preventing unprotected anal intercourse in gay men: A comparison of two intervention techniques. *International Journal of STD & AIDS*, 1995;**6**:89-94.

[17] Gold RS, Rosenthal DA. Examining self-justifications for unsafe sex as a technique of AIDS education: the importance of personal relevance. *International Journal of STD & AIDS* 1998;**9**:208-213.

[18] Guenther-Grey CA, Varnell S, Weiser JI, et al. Trends in Sexual Risk-Taking among Urban Young Men Who Have Sex with Men, 1999-2002. *Journal of the National Medical Association* 2005;**97**(7):38S–43S.

[19] Harding R, Bensley J, Corrigan N, Franks L, Stratman J, Waller Z, Warner J. Outcomes and lessons from a pilot RCT of a community-based HIV prevention multi-session group intervention for gay men. *AIDS Care* 2004;**16**:581–585.

[20] Hoff CC, Kegeles SM, Acree M, et al. Looking for men in all the wrong places. HIV prevention small-group programs do not reach high-risk gay men. *AIDS* 1997;**11**:829–31.

[21] Kelly JA, Murphy DA, Bahr R, Kalichman SC, Morgan MG, Stevenson Y, et al. Outcome of cognitive-behavioral and support group brief therapies for depressed, HIV-infected persons. *Am J Psychiatry* 1993;**150**:1679-1686.

[22] 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 population: an experimental analysis. *Am J Public Health* 1991;**8**:168-171.

[23] Miller RL. Assisting gay men to maintain safer sex: an evaluation of an AIDS service organization's safer sex maintenance program. *AIDS Educ Prevent* 1995;**7(suppl)**:48-63.

[24] Peterson JL, Coates TJ, Catania J, Hauck WW, Acree M, Daigle D, et al. Evaluation of an HIV risk reduction intervention among African-American homosexual and bisexual men. *AIDS* 1996;**10**:319-325.

[25] Roffman RA, Picciano JF, Ryan R, Beadnell B, Fisher D, Downey L, et al. HIV-prevention group counseling delivered by telephone: an efficacy trial with gay and bisexual men. *AIDS and Behavior* 1997;**1**:137-154.

[26] Shepherd J, Weare K, Turner G. Peer-led sexual health promotion with young gay and bisexual men: results of The HAPEER Project. *Health Education* 1997;**97**:204–212.

[27] Shoptaw S, Reback CJ, Peck JA, Yang X, Rotheram-Fuller E, Larkins S, et al. Behavioral treatment approaches for methamphetamine dependence and HIV-related sexual risk behaviors among urban gay and bisexual men. *Drug and Alcohol Dependence* 2005;**78**:125–34.

[28] Sorensen JL, Dilley J, London J, Okin RL, Delucchi KL, Phibbs CS. Case management for substance abusers with HIV/AIDS: a randomized clinical trial. *American Journal of Drug and Alcohol Abuse* 2003;**29**:133–150.

[29] Valdiserri RO, Lyter DW, Leviton LC, Callahan CM, Kingsley LA, Rinaldo CR. AIDS prevention in homosexual and bisexual men: results of a randomized trial evaluating two risk reduction interventions. *AIDS* 1989;**3**:21-26.

[30] Wilton L, Herbst JH, Coury-Doniger P, et al. Efficacy of an HIV/STI Prevention Intervention for Black Men Who Have Sex With Men: Findings from the *Many Men, Many Voices* (*3MV*) Project. *AIDS and Behavior* 2009;**13**:532-44.

[31] Wolitski RJ, Gómez CA, Parsons JT, and the SUMIT Study Group. Effects of a peer-led behavioral intervention to reduce HIV transmission and promote serostatus disclosure among HIV-seropositive gay and bisexual men. *AIDS* 2005;**19(Suppl 1)**:S99-S109.

[32] Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychol Meth* 1998;**3**:486-504.

[33] Gut A. Probability: A Graduate Course. 2005. Springer-Verlag.

[34] Senn SJ. Covariance analysis in generalized linear measurement error models. *Stat Med* 1990;**9**:583–6.

[35] Maris E. Covariance adjustment versus gain scores revisited. *Psychol Methods* 1998;**3**:309–27.

[36] Van Breukelen GJP. ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology* 2006;**59**:920–925.

[37] Whitehead A, Baily AJ, Elbourne D. Combining summaries of binary outcomes with those of continuous outcome in a meta-analysis. *Journal of Biopharmaceutical Statistics* 1999;**9**:1–16.

[38] Suissa S. Binary methods for continuous outcomes: a parametric alternative. *J Clin Epidemiol* 1991;**44**:241–248.

[39] Flanders WD, Kleinbaum DG. Basic models for disease occurrence in epidemiology. *International Journal of Epidemiology* 1995;**24**:1–7.

[40] Cheung YB. Zero-inflated models for regression analysis of count data: a study of growth and development. *Statistics in Medicine* 2002;**21**:1461–1469.

[41] Rose CE, Martin SW, Wannemuehler KA, Plikaytis BD. On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. *Journal of Biopharmaceutical Statistics* 2006;**16**:463-81.

[42] Severini TA. *Elements of Distribution Theory*, Cambridge University Press 2005, example 13.3.

Figure 1. Baseline Counts in 8 Studies
Counts for 8 separate studies truncated at 19
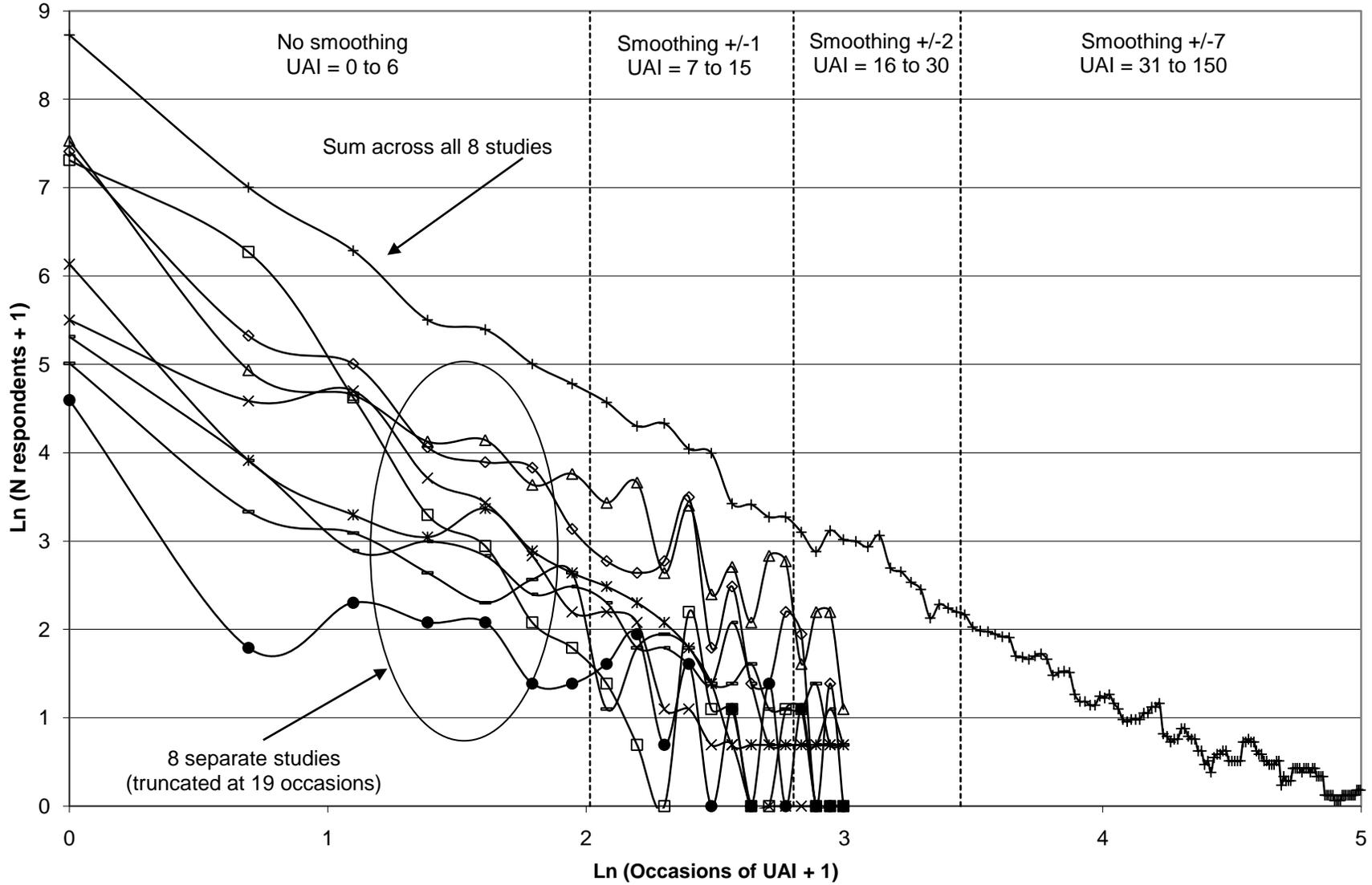
Figure 2. Logit of Proportion and Ln of Variance of Y by Ln of Mean
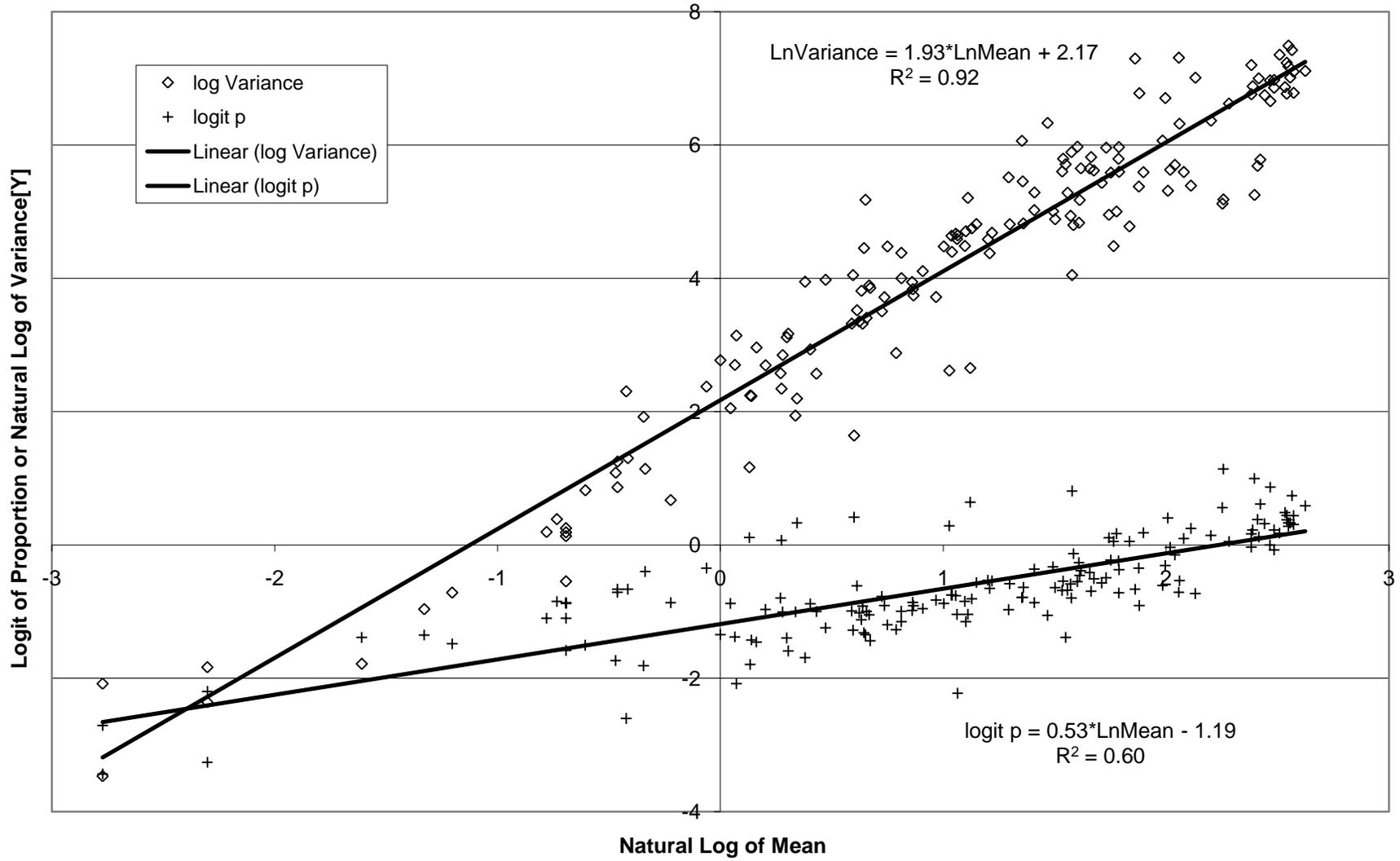
Figure 3. Observed Ln Rate Ratio and Predicted Ln Risk Ratio by Observed Ln Risk Ratio
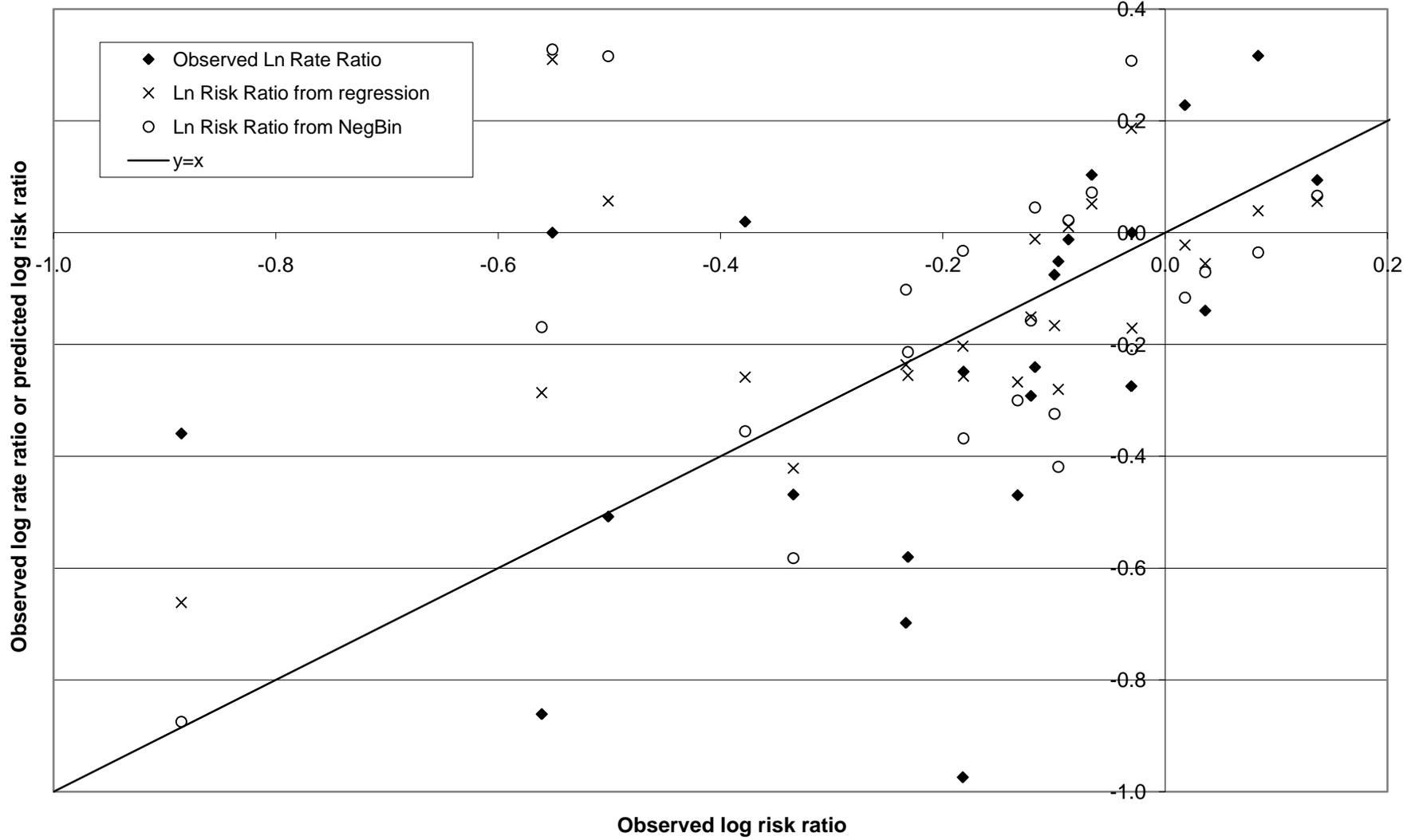
Figure 4. Ln Observed Weight Ln Rate Ratio and Predicted Weight Ln Risk Ratio by Ln Observed Weight Ln Risk Ratio
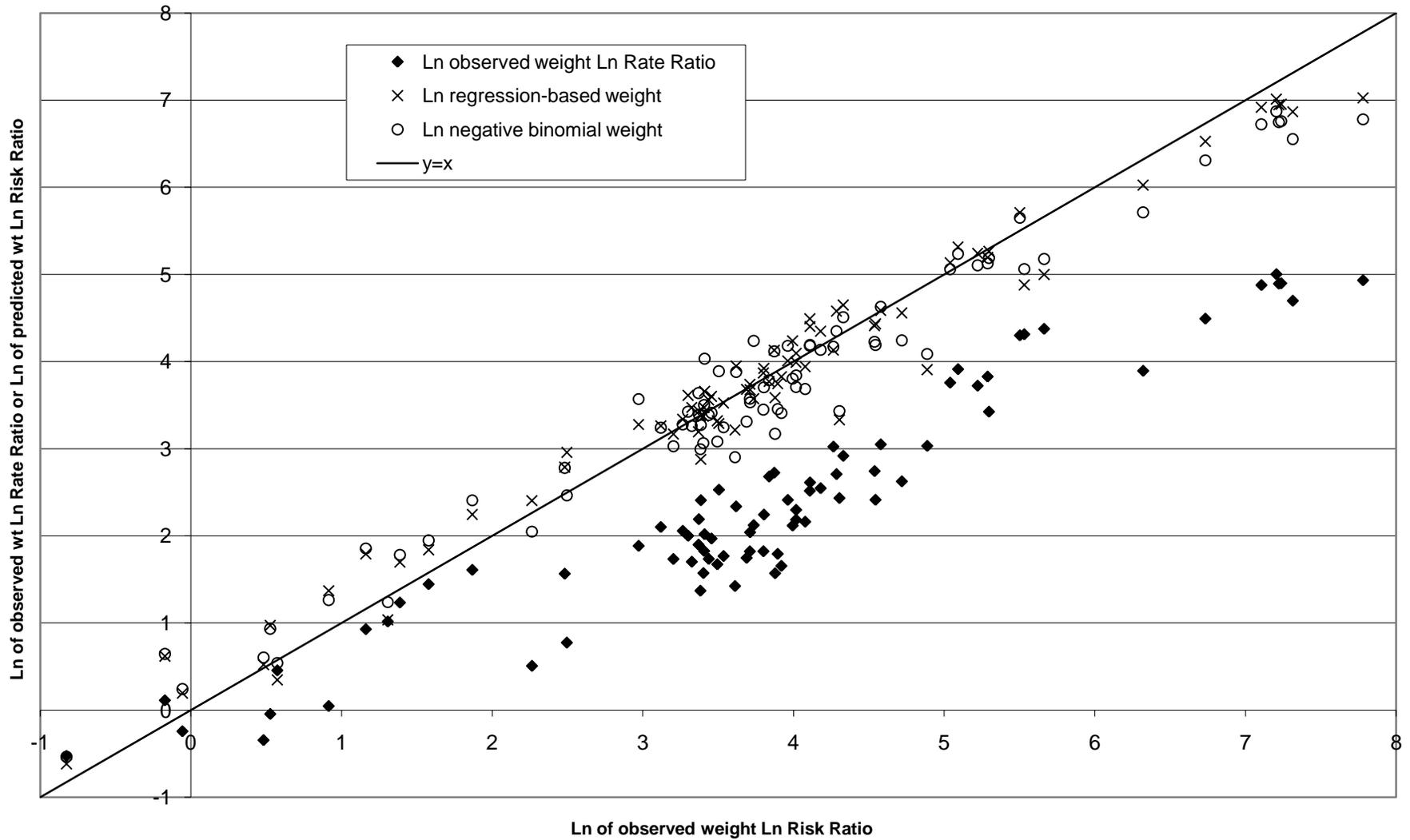
Figure 5. Observed Ln Risk Ratio and Predicted Ln Rate Ratio by Observed Ln Rate Ratio
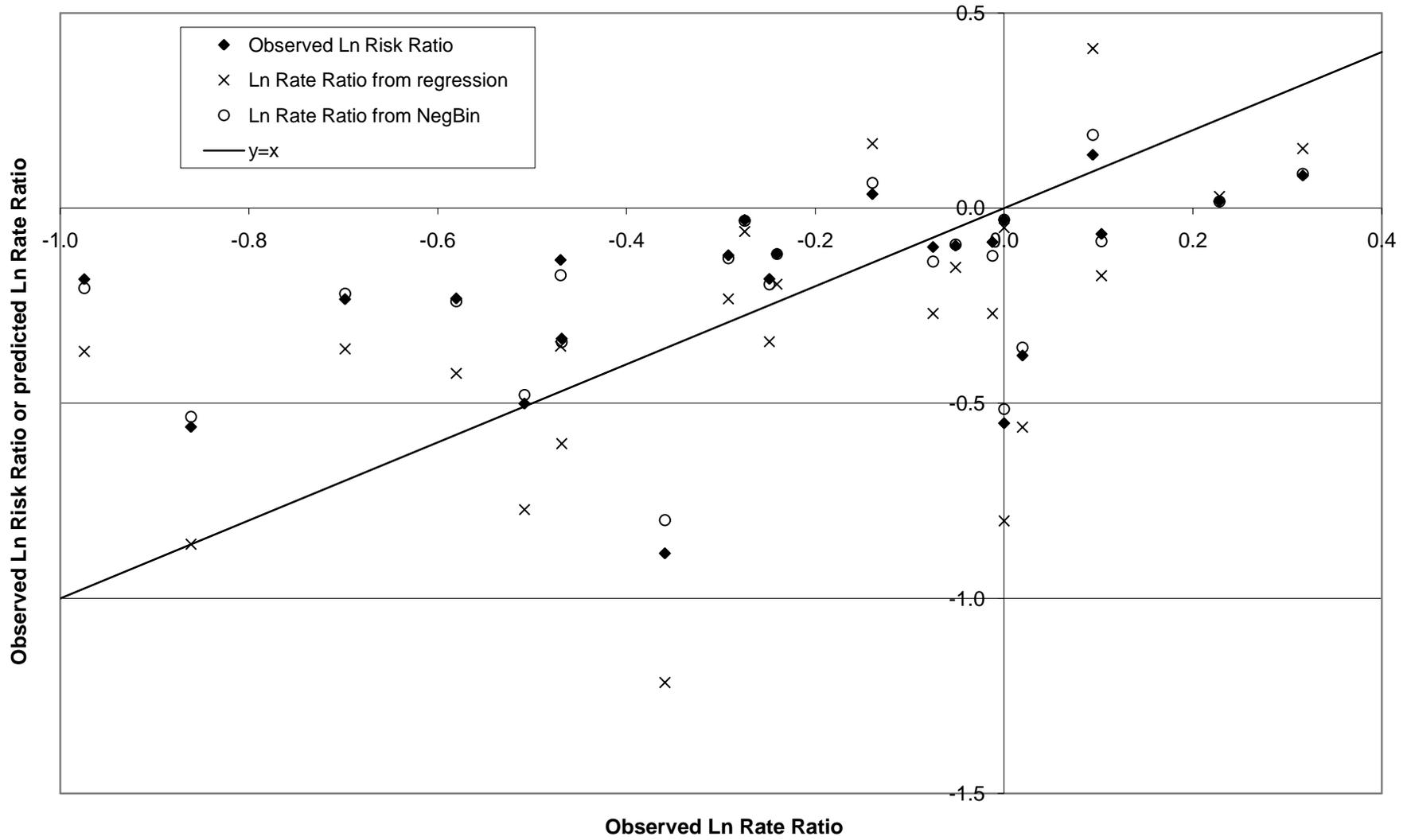
Figure 6. Ln Observed Weight Ln Risk Ratio and Predicted Weight Ln Rate Ratio by Ln Observed Weight Ln Rate Ratio
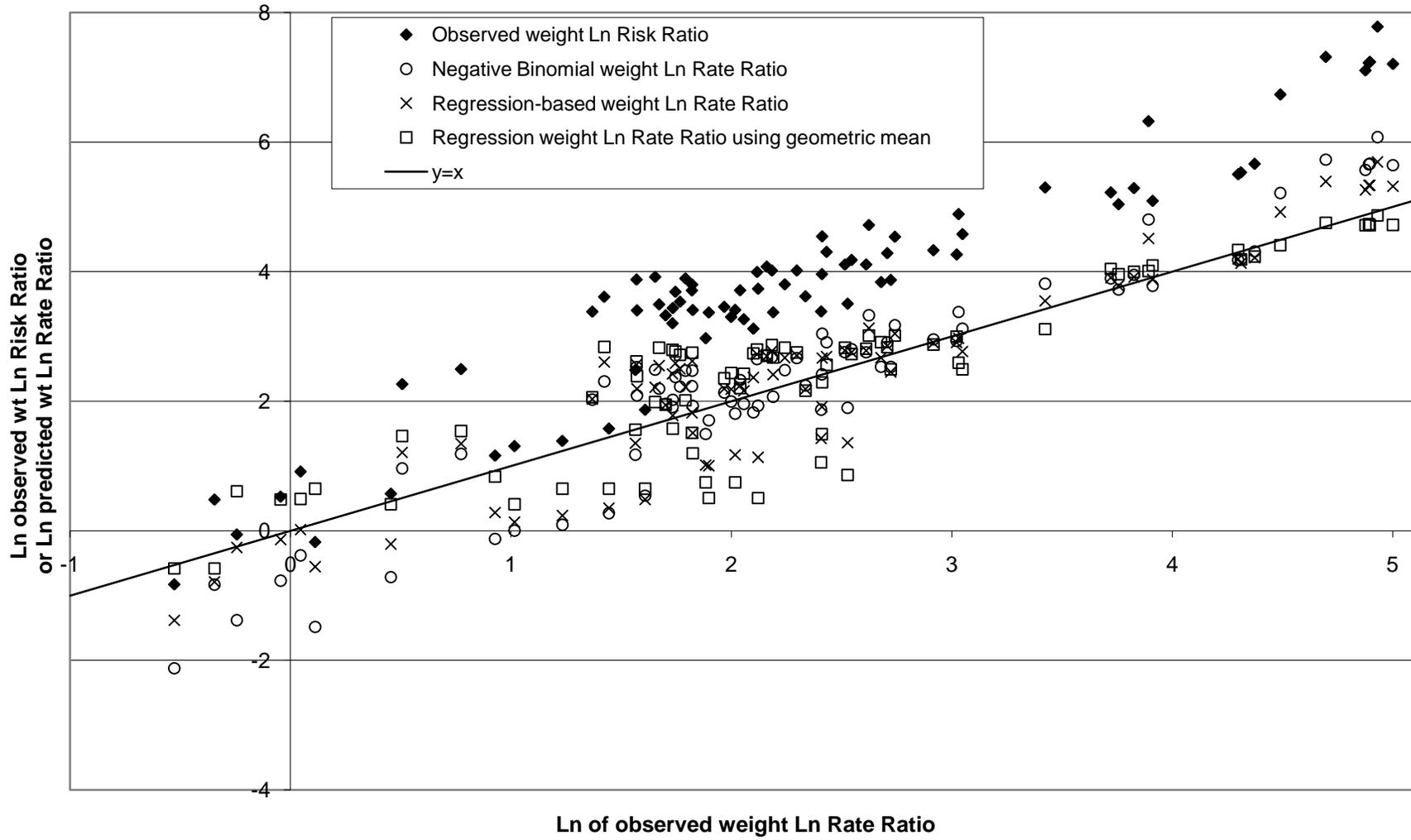
Figure 7. Ln Observed Weight Ln Rate Ratio and Ln Delta Estimates of Weight Ln Risk Ratio by Ln Observed Weight Ln Risk Ratio
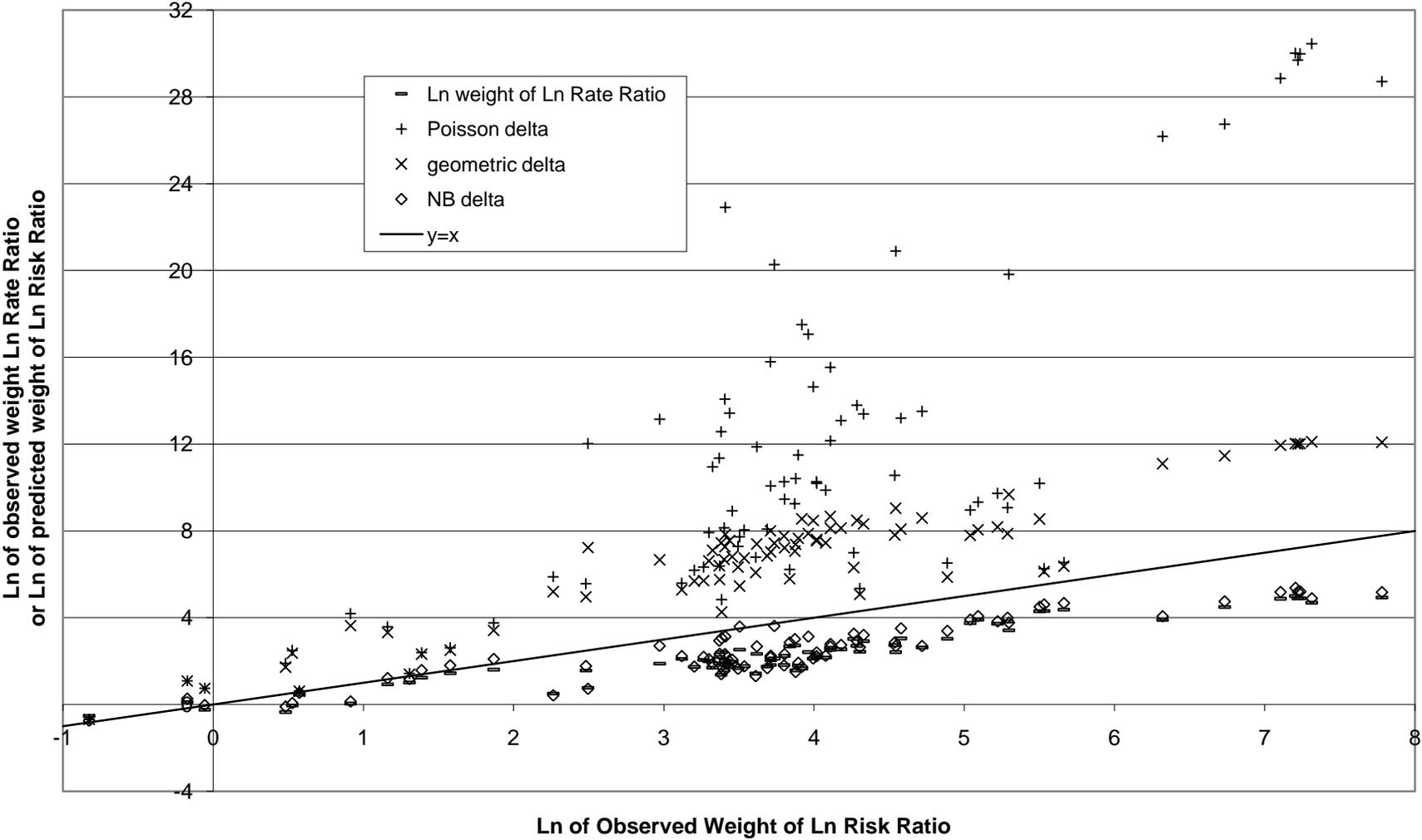
Figure 8. Ln Observed Weight Ln Risk Ratio and Ln Delta Estimates of Weight Ln Rate Ratio by Ln Observed Weight Ln Rate Ratio
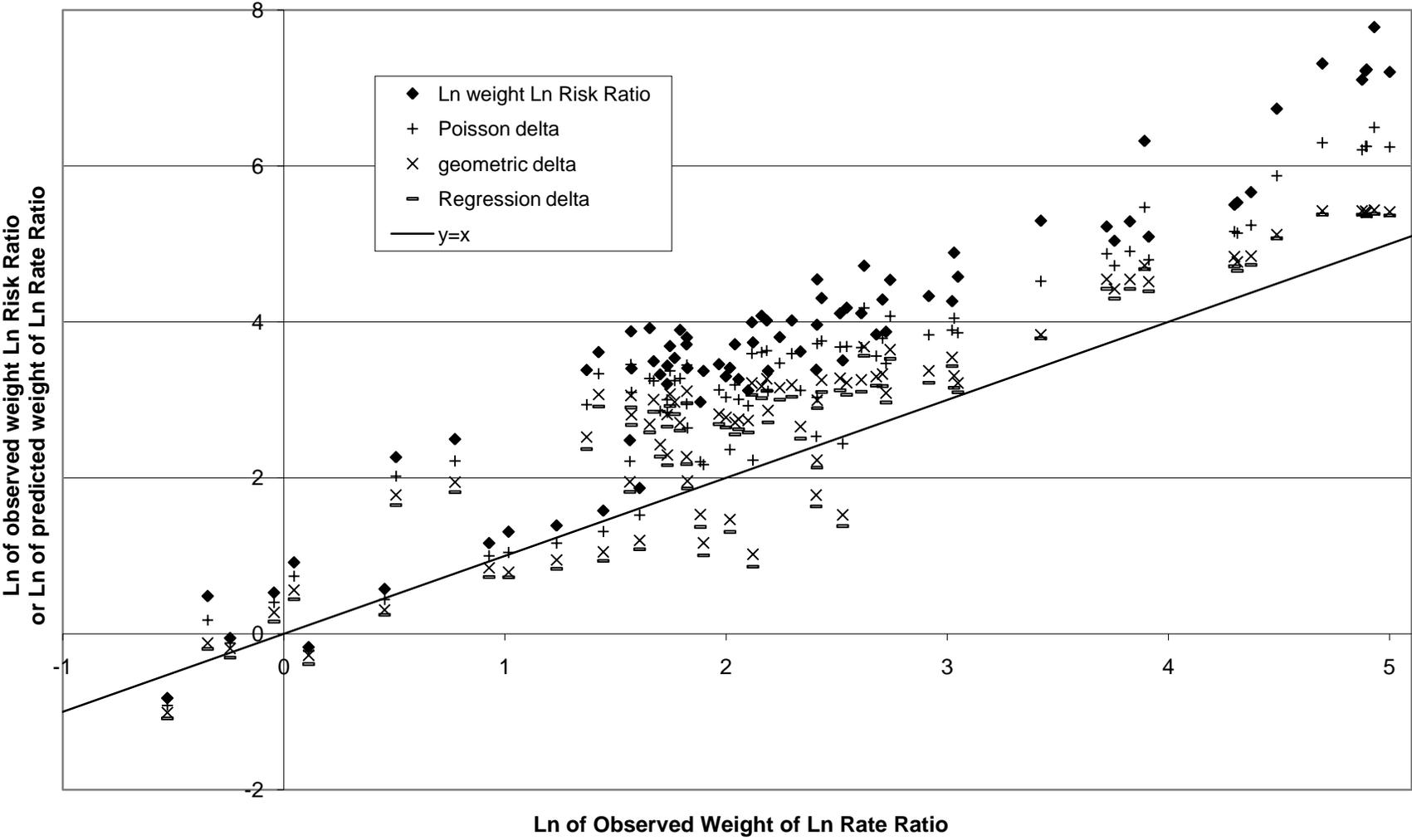
Appendix C. Meta-Analysis of HIV Prevention Research for Men Who Have Sex with Men

This published article is reproduced here with permission from *AIDS Education and Prevention*

HIV INTERVENTION RESEARCH FOR MEN WHO HAVE SEX WITH MEN: A 7–YEAR UPDATE

Wayne D. Johnson, David R. Holtgrave,
William M. McClellan, W. Dana Flanders,
Andrew N. Hill, and Michael Goodman

We conducted a systematic review and meta–analysis to locate, characterize, and summarize effects of behavioral HIV prevention interventions for men who have sex with men (MSM). We found 54 interventions with 16,224 participants that were evaluated in 40 randomized trials and controlled observational studies with independent comparison groups. Formats included 26 small group interventions, 18 individual–level interventions, and 10 community–level interventions. Fifteen interventions focused on HIV–positive individuals including MSM. The 38 interventions that were compared with minimal or no HIV prevention interventions, reduced unprotected sex by 27% (95% confidence interval [CI] = 15–37%). The other 16 interventions reduced unprotected sex by 17% beyond changes observed in standard or other HIV prevention interventions (CI = 5–27%). Behavioral interventions reduce self–reported unprotected sex among MSM.

Behavioral prevention remains central to the effort to reduce HIV transmission. Although antiretroviral therapy has tremendous lifesaving potential, it is expensive, does not cure, and can have debilitating side effects (Conant, 2004). Risk behaviors may increase if people believe that new treatments reduce subsequent transmission (Gray, et al., 2003). And an effective vaccine is still elusive (Garber, Silvestri, & Feinberg, 2004).

Men who have sex with men (MSM) still constitute the largest proportion of new infections in most of the developed world (Catania, 2000). In the 32 U.S. states that reported HIV infection from 2000 to 2003, the case rate increased among MSM (by 8% among blacks and Hispanics and 4% among Whites) while decreasing among high–risk heterosexuals and injection drug users (Centers for Disease Control and Prevention [CDC], 2004). Most new diagnoses of HIV/AIDS in the United States oc-

curred among people aged 25-34 (37%) or 35-44 (32%), with the rest divided about evenly between those under 25 (15%) and those 45 or older (16%). Man–to–man sex was by far the most common route of transmission among men in all race/ethnic groups.

Effects of behavioral interventions for MSM have recently been evaluated in numerous randomized trials and strong quasi–experimental studies. Quantitative synthesis can help to optimize the usefulness and interpretability of results across studies. Previous meta–analyses of HIV prevention interventions for MSM through 1997 found a 26% reduction in unprotected sex compared with neutral or standard conditions (Johnson, Hedges, & Diaz, 2003; Johnson, Hedges, et al., 2002). An update to that analysis found favorable effects among 17 studies that reported a basis in behavioral theory (odds ratio [OR] = .65; 95% confidence interval [CI] = .55, .77) and no effect among 3 that did not (OR = 1.03; 95% CI = .61–1.75) (Herbst et al., 2005).

In the present article we further update those meta–analyses to include 54 interventions evaluated with MSM in 40 studies. Our research questions were as follows: What behavioral interventions to reduce risk of HIV transmission among MSM have been tested in randomized trials or in rigorously controlled quasi–experimental studies? What populations have been served or underserved in these studies? What were the effects of these interventions? How do effects vary according to populations, content, and study design?

METHODS
SEARCH STRATEGIES AND ELIGIBILITY CRITERIA

We systematically reviewed the HIV prevention literature to find studies measuring the effects of behavioral interventions for MSM (Semaan et al., 2002). Resources included online databases (e.g., Medline, PsychInfo, PubMed, AIDSLine, Web of Science), reviews and other studies in the HIV prevention literature, expert recommendation, hand searches of selected journals, and manuscripts and unpublished reports submitted by researchers.

Key words for electronic searches varied according to database. As an example, a Medline search in August 2004 for <AIDS prevention & control [pc] or HIV infections/pc or sexually transmitted diseases/pc> yielded 24,143 citations. A search for <homosexuality or bisexuality or gay.mp or bisexual.mp or men who have sex with men.mp or seropositivity/psychology> yielded 13,262 citations, and a search for <randomization or intervention studies or program evaluation or random.mp or randomize.mp or randomized.mp or randomly.mp> yielded 292,874 citations. Most quasi–experimental studies included the terms *intervention studies* or *program evaluation*. Of the 77 citations included in all three searches, 49 were potentially eligible trials or reviews of HIV prevention interventions. Review of these 49 led to identification of 21 trials that were eligible by the criteria described below.

The potentially eligible HIV prevention studies from all sources were then evaluated by criteria of outcomes measured and study design. We included only studies that measured intervention effects on behaviors understood to affect risk of HIV transmission (e.g., unprotected sex, condom use, number of partners) and biological outcomes, including incidence of infection by HIV or other sexually transmitted diseases (STD). We defined *unprotected sex* as anal intercourse without a condom. Data concerning other sexual and drug use behaviors were not frequently available. Only three eligible studies reported biological outcomes; we will consider these in the discussion section. Because unprotected anal sex is the most epidemiologically pertinent behav-

ior for MSM (O'Leary, DiClemente, & Aral, 1997) and was available for all studies, we restricted our analyses to this outcome.

We excluded interventions that focused not on sexual transmission but on cognitive or affective outcomes such as distress associated with HIV testing, or health and coping for seropositive men (Chesney, Chambers, Taylor, Johnson, & Folkman, 2003; Perry, Fishman, Jacobsberg, Young, & Frances, 1991). We included only studies in which MSM constituted all or a substantial proportion of the study sample (e.g., HIV–seropositive individuals) or were specifically targeted by the intervention. When other populations were included, we obtained outcome data for the MSM subset or reduced the study weight to reflect only the proportion who were MSM.

Acceptable study designs were randomized controlled trials and certain quasi–experimental designs. Quasi–experimental studies were required to include independent comparison groups assigned without bias, that is, without regard to volition, self–selection, need, or other baseline characteristics, and to include separate baseline data for the intervention and comparison groups. We requested supplemental information from authors when separate results by study arm were not published.

CALCULATION OF EFFECT SIZES

Because eligible studies used randomized or quasi–experimental designs, we chose rate ratios (RR) to estimate intervention effects for count measures and prevalence ratios (PR) for dichotomous measures (Deeks, 1999; Greenland, 1998). For each study that reported count measures (number of episodes of or partners for unprotected sex), the rate ratio at follow–up was the ratio of the mean in the intervention group to the mean number in the comparison group. The natural logarithm of the rate ratio (LnRR) was then an estimate of the intervention effect. The reciprocal of the variance of LnRR (see Appendix) served as a measure of the weight of information provided by the study.

Similarly, the prevalence ratio at follow–up was the ratio of the proportion of respondents reporting unprotected sex in the intervention group to the proportion in the comparison group. The natural logarithm of the prevalence ratio (LnPR) was an estimate of the intervention effect, and the reciprocal of the variance of LnPR (see Appendix) estimated the weight of the study. Rate ratios and prevalence ratios less than one represented a difference favoring the experimental intervention group.

When individual–level data were available, we used SAS Proc Genmod to estimate intervention effects adjusted for the baseline value of covariates such as the outcome variable, age, race/ethnicity, and serostatus. For count outcomes, we used the negative binomial distribution and the log link function and adjusted the scale for Pearson's chi–square divided by its degrees of freedom to estimate the rate ratio. For dichotomous outcomes, we used the binomial distribution and the log link function to estimate the prevalence ratio.

When individual–level data or adjusted statistics were not available, we adjusted for the baseline distribution of the outcome variable by subtracting the baseline LnRR or LnPR from the follow–up LnRR or LnPR. We used the lesser of the baseline and follow–up weights for such studies. These decisions concerning baseline adjustment and weights were based on the available empirical examples with raw data.

In studies where communities are the unit of assignment to treatment, the variance of the intervention effect will be underestimated if intraclass correlation (ICC) is not accounted for (Murray, 1998). We derived the adjustment factor to reduce study weights where necessary to account for ICC. For small values of ICC, the adjustment

factor is approximately equal to Donner's variance inflation factor (VIF): $1 + ICC \times (m - 1)$ where m is the number of subjects in each unit of assignment. Derivation of this factor is available from the authors on request. We assumed an ICC of .005, the value observed in the one study for which ICC was published (Kelly et al., 1997).

For studies that measured results at multiple follow–up times, we used data representing cumulative effects closest to 12 months after the intervention. We used outcome variables that did not distinguish between insertive and receptive sex, main and nonmain partners, or partners perceived to be seroconcordant versus serodiscordant when such data were available. For studies from which the only available results were separated by insertive versus receptive sex, or main versus nonmain partners, we used the average point estimate and the average weight of the two measures to estimate the underlying combined effect (Johnson, Semaan, et al., 2002). We accepted results concerning only nonmain, serodiscordant, or unknown serostatus partners when results were not available concerning main or seroconcordant partners. For studies that compared two or more experimental interventions against a single control group, we divided the control group into equal parts for comparison to each of the interventions.

STATISTICAL ANALYSIS

We applied the standard procedures for meta–analysis to conduct summary, stratified, and regression analyses (Cooper & Hedges, 1994). We conducted separate analyses using rate ratios and prevalence ratios. To include all eligible studies in each analysis, we substituted prevalence ratios for rate ratios in studies that measured only dichotomous outcomes and vice versa in studies that measured only count outcomes. If the prevalence ratio is constant across cutpoints then the rate ratio equals the prevalence ratio. This assumption of constancy across cutpoints is analogous to the assumption used to justify transformation between log odds ratios and standardized mean differences in other meta–analyses (Chinn, 2000; Hasselblad & Hedges, 1995; Johnson, Semaan, et al., 2002). This assumption appeared plausible based on the studies for which effects at multiple cutpoints were available. Variances (and therefore weights) however differ substantially between count and dichotomous outcomes. We used the method of moments to develop an estimate of the variance of LnRR when only dichotomous data were available, and the variance of LnPR when only count data were available (see Appendix).

We present results separately for interventions contrasted against minimal to no HIV prevention control conditions and those contrasted against standard or other HIV prevention conditions. We defined minimal to no HIV comparison conditions as including no treatment, wait lists, lagged designs, counseling for emergencies only, passive display of materials in community settings, and several treatments not addressing sexual behavior (diet and exercise training, substance abuse treatment, health support groups, and medication adherence consultation). Standard or other comparison conditions included HIV prevention seminars, individual HIV prevention counseling and testing, HIV prevention videos, and keeping a diary of sexual activity in the context of HIV prevention.

We considered both random and fixed–effects meta–analyses (Hedges & Vevea, 1998). Because intervention effects were generally homogeneous, results of the two types of models were usually identical (Hedges, 1994). When results did differ, those from the fixed effects models were slightly (about 1%) more conservative. Therefore we present only results of fixed effects models. We used stratified analyses to examine subgroup effects according to intervention format (small group, individual, or com-

TABLE 1. Twenty–eight[a] Studies of Men Who Have Sex With Men (MSM), Comparing 38 HIV Prevention Interventions With Minimal or No HIV Prevention Control Conditions (by intervention format and in approximate order of dates conducted)

Study	Location and Year Conducted	Intervention(s) and Specified Populations	Comparison Condition
Small Group Interventions			
Kelly et al., 1989	Mississippi, 1987	AIDS risk education, cognitive behavioral self–management training, sexual assertion training, development of relationship skills and social support. 12 weekly 75–90–minute meetings.	Wait list
Coates et al., 1989	San Francisco, 1987	Stress management training for HIV–positive men, including systematic relaxation, health behavior change, and stress management skills. 8 sessions of 2 hours and 1 all–day retreat.	Wait list
Rosser et al., 1990*	Auckland, 1987–1988	(E) group discussion on eroticizing safer sex. (S) group discussion of safer sex guidelines and effect of AIDS. 2–2.5 hours.	No treatment
Tudiver et al., 1992	Toronto, 1990	(P1) Discussion of safer sex, personal experiences, coping strategies, skills, and role plays. One 3–hour session led by peer volunteer. (C4) Similar content, 4 weekly 2–hour sessions led by paid counselors.	Wait list
Roffman et al., 1998	Seattle, 1989–1991	Relapse prevention: HIV education, motivation, listening, self–talk, assertiveness, avoiding risky situations, debriefing, maintenance strategies, social support, self–esteem. 17 weekly sessions, 2 hours each.	Wait list
Peterson et al., 1996	San Francisco, 1989–1991	(3s) 3 weekly 3–hr sessions on AIDS risk education, cognitive behavioral self–management training, assertion training, and attempts to develop self–identity and social support for African–American MSM. (1s) Single 3–hr session of same content.	Wait list
Kelly et al., 1993	Milwaukee, 1991	Depressed HIV–positive men were assigned to either (CB) cognitive behavioral approach with behavioral or skill training themes, (SS) social support group, or control. 8 weekly 90–minute sessions (each arm).	Individual crisis therapy on request
Stall et al., 1999	San Francisco, 1990–1993	Closed group treatment for substance use disorder plus exercises concerning sexual risk taking, for MSM attending a nonresidential treatment center. 16–week program, two 3–hour sessions per week.	Group treatment for substance use only
Roffman et al., 1997	Western US, 1992–1994	Small–group telephone conferences of geographically dispersed MSM; permitted anonymity and participation by men in rural areas. Coping strategies to deal with high–risk situations, setting realistic, client–centered risk reduction goals, identifying antecedents to risk behavior. 14 weekly 90–minute phone calls.	Wait list
Choi et al., 1996	San Francisco, 1992–1994	Culturally specific group counseling for Asian and Pacific Islander MSM, including development of positive self–identity and social support, safer sex education, eroticizing and negotiating safer sex. One 3–hr session.	Wait list
Rotheram–Borus et al., 2001	Los Angeles, New York, San Francisco, Miami, 1994–1997	Youth with HIV at 4 clinical care sites. 23 weekly 2–hr sessions in two modules. Staying Healthy, (e.g., coping with learning HIV status, disclosure, health care decisions) and Acting Safe (e.g., protecting self and partner, safer sex options, drugs and alcohol, avoiding internal and external triggers, anxiety and anger). 63% of the youth were MSM.	Usual clinical activities

572

Study	Location, Date	Description	Comparison
Kalichman et al., 2001	Atlanta, 1997	Support group for people with HIV, most of whom were MSM, 74% African Americans. Create sexual health and relationship plans, develop communication and disclosure skills, learn hazards of coinfection with other STIs. Five 120–minute sessions both for intervention and control.	Support group for health maintenance
Harding et al., 2004	London, 2000	"SM Sex: An Introduction to the SM scene." Sessions address assumptions and knowledge, practical tools of sadomasochistic sex, risk taking, emotional aspects, sexually transmitted infections and HIV transmission, rights and responsibilities, legal issues, the role of fantasy, and limits and boundaries. Up to 25 group members, 4 sessions of 7 hours.	Wait list
Carballo–Diéguez et al., 2004	New York, 1998–2002	Latino MSM. 8 sessions on themes of oppression, transgression of rules, excuses (or rationalizations), substance use, goal setting, the role of pleasure, self–efficacy and plans for the future. Exercises included word association, story analysis, problem solving, analysis of Spanish dichos (proverbs), discussion of participants weekly sexual diaries.	Wait list

Individual–Level Interventions

Study	Location, Date	Description	Comparison
Rosser et al., 1990*	Auckland, 1987–1988	(V) 15–minute safer sex video. (C) 30–minute individual counseling using a behavioral HIV risk assessment system.	No treatment
Picciano et al., 2001	Seattle, 1998–1999	MSM who report 3 or more recent episodes of oral or anal sex without condoms were given feedback by telephone regarding a baseline risk assessment. 1 hour.	Wait list
Patterson et al., 2003	San Diego, 1999–2001	HIV–positive volunteers (85% gay or bisexual) recruited by posters, service providers and others, and reporting unprotected sex with HIV–negative or unknown status partners, received social cognitive theory–based intervention in 1 or more of 3 domains (condom use, negotiation of safer sex, disclosure of HIV status) in 1 of 3 formats: (T) 90–minute targeted (participant selected domain[s] of concern) (C) 90–minute comprehensive (all 3 domains), or (B) 90–minute comprehensive plus two 90–minute booster sessions.	Three 90–minute sessions on diet and exercise
Richardson et al., 2004	California, 1999	Prevention counseling from medical providers supplemented with written information. Two clinics used a gain–framed approach (G) (positive consequences of safer–sex) and two used a loss–frame (L) approach (negative consequences of unsafe sex). 3–5–minute every visit; 4 hours training for clinicians.	2 attention–control clinics (medication adherence)
Rotheram–Borus et al., 2004	Los Angeles, New York, San Francisco, 1999–2002	Modification of RB 2001 to individual format to increase participation. Substance–using youth with HIV, most referred by social service agencies or medical providers. 18 weekly 2–hr sessions. Improving physical health, maintaining drug regimens, coping with learning HIV status, health care decisions. Reducing unprotected sex and substance use, examining trigger situations, condom use and negotiation skills and self–efficacy. Focus on condom use rather than disclosure. Reducing distress, anticipating situations that raise anxiety, depression, fear, or anger. Recognizing and controlling negative emotion with relaxation, self–instruction, meditation. Identifying life goals. 69% of the youth were MSM. Delivered by telephone (T) or in person (IP).	Wait list

Community–Level Interventions

Study	Location, Date	Description	Comparison
Kelly et al., 1991	Mississippi, Louisiana, 1989	Popular opinion leaders were trained to serve as behavior change endorsers to their peers in gay clubs. 4 weekly 90–minute training sessions for opinion leaders.	Lagged design
Kelly et al., 1997	4 U.S. states, 1991–1994	Popular men in gay bars were engaged to advocate benefits of behavior change to peers, and HIV educational materials were placed in bars. West Virginia, Washington, New York, Wisconsin; 5 weekly 2–hr training sessions for opinion leaders.	Educational materials in bars only

TABLE 1. (continued)

Study	Location and Year Conducted	Intervention(s) and Specified Populations	Comparison Condition
Kegeles et al., 1996	Oregon, California, 1992–1994	8–month peer–led program for young gay men including outreach, small groups, and a publicity campaign.	Lagged design
CDC ACDP, 1999	3 Western U.S. cities, 1991–1995	32–month intervention for MSM who do not self–identify as gay, surveyed in public sex environments. Distribution and discussion of flyers containing condoms and role–model stories from men in the community about making progress toward consistent condom use. Seattle; Denver; Long Beach, CA..	No treatment
Hoff et al., 1997	2 Western U.S. cities, 1994	The 18–month intervention (Portland, OR) targeted community mobilization, social support, education, outreach, volunteer coordination, HIV testing, and provider mobilization.	No treatment (Tucson, AZ)
Shepherd et al., 1997	Southampton (England), 1996	The HAPEER Project. Young peer educators administered a structured interview to peers (mean age 24 years) in a range of environments to recruit for study, collect baseline data, initiate discussion of sexual health, and identify and respond to individual sexual health needs.	A neighboring gay community
Miller et al., 1998	New York, 1995–96	Replication of Kelly (1991) in hustler bars. 57% of respondents self–identified as gay, 31% as bisexual. 3 training sessions of 2 hours in 1 wk for opinion leaders.	Lagged design
Elford et al., 2001	London, 1997–1998	Replication of Kelly (1991) in 5 gyms. Difficulties reported in delivering the intervention; only 3% of respondents reported having been spoken to by volunteers. First assessment at 6 months; ongoing for 18 months.	Lagged design
Flowers et al., 2002	Glasgow, Edinburgh, 1996–1999	Delivered through gay bars in Glasgow. Peer–led sex health promotion, gay–specific genito-urinary medicine services, free–phone hotline with sex health information and details of local sexual health services, endorsement of testing, risk assessment, and sexual health. 9 months in community, 2 days training for peer educators.	No treatment
Kegeles et al., 2002	3 Southwestern U.S. cities, 1997–1998	Based on theories of empowerment, diffusion and peer mobilization, the 12–month intervention (Albuquerque, NM) featured a young gay men's community center, a core group of men who ran the project, informal outreach among friends, formal outreach at gay venues and social events, and small groups focused on safer sex and informal outreach.	No treatment (Austin, TX; Phoenix, AZ)

Note. STI = sexually transmitted infections. Other abbreviations, which appear in parentheses, correspond to suffixes in Figure 1. [a]One study (Rosser 1990) included both small–group interventions and individual–level interventions.

574

munity level), and to summarize interventions for HIV–positive MSM. We applied the standard principles of weighted meta–regressions (Cooper & Hedges, 1994) to account for multiple study characteristics and to examine differences in effects according to exposure rates in community–level interventions.

We examined the potential effect of outliers by excluding each intervention effect one at a time and recalculating the summary effect. To investigate the possibility of publication bias, we examined a linear regression through the funnel plot of treatment effect on sample size (Macaskill, Walter, & Irwig, 2001). To be concise, we present regressions, sensitivity to outliers, and analysis of publication bias only for rate ratios and not prevalence ratios.

RESULTS

As of May 2005, we had identified 54 experimental HIV prevention interventions for MSM evaluated in 40 eligible studies. Primary citations for these studies were found in 19 journals and one conference, with the largest numbers published in *AIDS* (8 studies), and the *American Journal of Public Health*, (5 studies). Eleven studies tested two or more experimental interventions against comparison conditions. We treat the total of $k = 54$ experimental interventions and their associated control data as separate units for description and analyses.

Most of the interventions ($k = 38$) were compared with minimal or no HIV prevention control conditions (Table 1). Of these, 18 interventions were delivered in small-group format, 10 in individual–level format, and 10 in community–level format. The other 16 interventions, including 8 small-group interventions and 8 individual–level interventions, were compared with standard or other HIV prevention conditions (Table 2).

Over two–thirds of the 54 interventions ($k = 38$) were evaluated in the United States. Also represented were England ($k = 4$), Australia ($k = 4$), New Zealand ($k = 4$), Canada ($k = 2$), Scotland ($k = 1$), and Brazil ($k =1$). The weighted mean age of participants was 34 years (range 21 to 42).

Across all studies, about 31% of study participants were African American, Latino, Asian, or of other race/ethnic groups besides Whites. Only four interventions in three studies focused on specific racial or ethnic groups: African Americans (Peterson et al., 1996), Asians and Pacific Islanders (Choi, et al., 1996), and Latinos (Carballo–Dieguez et al 2005). In five U.S. studies representing seven interventions (Cleary et al., 1995; EXPLORE Study Team, 2004; Kalichman, et al., 2001; Rotheram–Borus et al., 2001; Wolitski, Parsons, Gómez, & the SUMIT Study Group, 2005), 49%-81% were African American or Latino. In the study in Brazil (Sampaio, Brites, Stall, Hudes, & Hearst, 2002), 51% described themselves as Mulatto, 34% as White, and 15% as Black. In four more U.S. studies (Kelly et al., 1993; Miller, Klotz, & Eckholdt, 1998; CDC AIDS Community Demonstration Projects Research Group, 1999; Patterson, Shaw, & Semple, 2003) evaluating seven interventions, more than a third of participants were ethnic minorities.

Ten studies, including 15 interventions, focused on HIV–positive populations. In three of these studies (Coates, McKusick, Kuno, & Stites, 1989; Kelly, et al., 1993; Wolitski et al., 2005), all or nearly all (94% to 100%) participants were MSM. In the other studies of HIV–positive individuals (Cleary et al., 1995; Kalichman et al., 2001; Patterson et al., 2003; Richardson et al., 2004; Rotheram–Borus et al., 2001, 2004; Sorensen et al, 2003), the majority (55%-80%) of participants were MSM. HIV prevalence was particularly high in three other studies: Shoptaw et al. (2004; 61%), Stall,

TABLE 2. Twelve Studies of Men Who Have Sex with Men (MSM), Comparing 16 HIV Prevention Interventions to Standard or Other HIV Prevention Control Conditions (by intervention format and in approximate order of dates conducted)

Study[a]	Location and Year Conducted	Intervention(s) and Specified Populations	Comparison Condition
Small–Group Interventions			
Valdiserri et al., 1989	Pittsburgh, 1986–1987	AIDS information and safer sex lecture (60–90 minutes) followed by skills training, discussion and rehearsal of safer sex negotiation.140 minutes.	Lecture only
Cleary et al., 1995	New York, 1986–1988	Blood donors testing HIV–positive received individual counseling (IC) plus a cognitive behavioral and skills training support group to provide more detailed information, encourage risk reduction behavior, provide support, and facilitate functional coping responses. 6 weekly meetings of 90 minutes.	IC plus community referral
Imrie et al., 2001	London, 1995–1998	Cognitive behavioral intervention at sexual health clinic for gay men with acute STI or unprotected sex in past year. Standard management (20 minute, one–on–one counseling and referrals) plus 1–day small–group workshop.	Standard management only
Rosser et al., 2002	Minneapolis, 1997–1998	Comprehensive seminar featuring systematic desensitization to sexual education curriculum, study of homosexual identity formation, sexual health education, research on cofactors of unsafe sex (drugs, alcohol, loneliness, falling in love). 2 full–day sessions.	3–hour HIV prevention video
Sampaio et al., 2002	Salvador, Brazil, 1998–1999	Safer sex workshop with games, role–playing, small group discussion using verbal and nonverbal communication. Basic information, clarification of misconceptions, recognition of risk. Nongenital practices, safe sex in committed versus other relationships, mechanics of using condoms, strategies for refusing unsafe sex, negotiating new sexual patterns. One 3–4–hour session with 15 to 20 subjects.	1–hour lecture & discussion infectious disease, condom skills
Shoptaw et al., 2005	Los Angeles, 1998–2000	Four 16–week alternatives for meth–dependent MSM. CBT: Group education to initiate meth abstinence and quickly resume abstinence if relapse occurs. Internal and external triggers, stages of recovery from meth dependence, identification of emotional states signaling relapse. Cognitive skills such as thought stopping, craving management, relapse analysis, adoption of healthy lifestyle behaviors. 90 minutes 3 times a week. CM: Vouchers (e.g., for groceries, camera equipment, plane fare to visit family, clothing) for drug abstinence. Escalating value for successive negative urine samples with reset after relapse; max total value $1,300. CBT + CM: both treatments simultaneously. GCBT: enhancement of CBT to address gay–specific context. Difference in sexual behavior on and off drug, indicators of meth use in sexual partners and friends, comparison of revealing one's drug problem to the coming out process, examples from circuit parties. We analyzed these as pair "g" (GCBT vs. CBT) and pair "c" (CBT+CM vs. CM only).	Pair 1: CBT only. Pair 2: CM only
Wolitski et al., 2005	New York, San Francisco 2000–2002	Group activities facilitated by HIV–positive peers. Building community for HIV–positive individuals, information, personal responsibility, assumptions and disclosure of serostatus, communication skills, effects of substance use and behavior on immune system, coping with HIV, mental health. Six 3–hour sessions.	90 minute forum and lecture

Individual–level interventions

Gold et al., 1995	Melbourne, Sydney, 1992–1993	Men who had recently had anal sex without a condom were assigned to keep a diary of sexual activity for 16 weeks and either (SJ) provide self–justifications for 2 episodes of anal sex without a condom, or (PO) examine and describe 10 AIDS education posters.	Diary only
Gold et al., 1998	Melbourne, Sydney, 1995–1996	Men who had recently had anal sex without a condom were assigned to keep a diary of sexual activity for 16 weeks and either (SJP) examine and describe 10 AIDS education posters highlighting the pitfalls of self–justifications or (RSE) describe in detail a recent sexual encounter that included anal sex without a condom.	Diary only
Sorensen et al., 2003	San Francisco, 1994–1997	12 months intensive case management for substance abusers with HIV/AIDS. Hybrid of full service and referral models. Case managers were paraprofessionals, former consumers of HIV, or substance abuse treatment services. Subjects were recruited at a teaching and public hospital from inpatient medical wards, outpatient heroin detox clinic, and emergency department. 48% were MSM.	Brief contact: education, risk reduction information, referrals
Dilley et al., 2002	San Francisco, 1997–2000	2 pairs of intervention and comparison conditions for MSM attending clinic for HIV standard counseling and testing (CT). No–diary intervention (N): CT + self–justifications session (1 hr), where the client reviewed and challenged his own self–justifications for a recent occasion of unsafe sex. Diary intervention (D): CT, self–justifications, and 90–day sexual diary.	N: CT only. D: CT plus diary
Explore et al., 2004	6 US cities, 1999–2003	10 one–on–one counseling sessions followed by maintenance sessions every 3 months. Risk assessment, sexual communication, knowledge of HIV serostatus, alcohol and drug use, triggers for unsafe sex, motivational interviewing. Total span up to 48 months.	Twice–yearly counseling and HIV testing

Note. STI = sexually transmitted infection; CBT = cognitive behavioral therapy; CM = contingency management; meth = methamphetamine. Other abbreviations, which appear in parentheses, correspond to suffixes in Figure 2.

Paul, Barrett, Crosby, and Bein (1999; 50%), and Carballo–Diéguez et al. (2005; 36%). Among the remaining 15 interventions for which HIV prevalence was reported and HIV status was not an inclusion or exclusion criterion, the weighted prevalence of HIV was 14%.

EFFECTS OF INTERVENTIONS VERSUS MINIMAL
TO NO HIV PREVENTION

The 38 interventions that were contrasted against minimal to no HIV prevention comparison conditions reduced unprotected sex by 27% (95% CI = 15%, 37%) (Figure 1). The corresponding rate ratio was .73 (CI = .63, .85). This effect represents a decrease from the average background mean of 10.1 unprotected episodes in a 6–month period to 7.4 (CI = 6.3, 8.6), and from 1.2 partners for unprotected sex in a 6–month period to 0.9 (CI = 0.8, 1.0). The intervention effects were statistically homogeneous (Q_{37} = 30.5, p = .76). In subgroup analyses the rate ratio was .71 (CI = .57, .89) among 18 small-group interventions, .87 (CI = .60, 1.26) among 10 individual–level interventions, and .70 (CI = .54, .90) among 10 community–level interventions.

The same 38 interventions reduced the proportion of subjects reporting unprotected sex by 16% (CI = 10%, 21%). The corresponding prevalence ratio was .84 (CI = .79, .90). This effect represents a decrease from an average of 41% reporting unprotected sex to 35% (CI = 32%, 37%). In subgroup analyses, significant favorable effects were observed for small-group (PR = .80; CI = .72, .89) and community interventions (PR = .86; CI = .76, .96). Effects among individual–level interventions were also favorable (PR = .93) but not statistically significant (CI = .78, 1.10).

EFFECTS OF INTERVENTIONS VERSUS STANDARD
OR OTHER HIV PREVENTION

The 16 remaining interventions reduced unprotected sex by 17% beyond changes observed in standard or other HIV prevention interventions (RR = .83; CI = .73, .95) (Figure 2). In subset analyses, rate ratios were .75 (CI = .60, .93) among eight small group interventions and .88 (CI = .75, 1.04) among individual–level interventions. There were no community–level interventions in this subset.

The same 16 interventions reduced the proportion reporting unprotected sex by 6% beyond changes observed in standard or other HIV prevention interventions (CI = 2%, 10%). The corresponding prevalence ratio was .94 (CI = .90, .98). The reduction was 10% among the eight small-group interventions (PR = .90; CI = .83, .99), and 5% among the eight individual–level interventions (PR = .95; CI = .91, 1.00).

INTERVENTIONS FOR HIV–POSITIVE MSM

A 21% reduction in unprotected sex (RR = .79; CI = .61, 1.02) was observed among the 15 interventions for HIV–positive individuals (total MSM sample size = 2,164). Effects were more clearly favorable among the seven small-group interventions (RR = .71; CI = .51, .99), two of which were contrasted against other HIV interventions, than among the eight individual–level interventions (RR = .91; CI = .62, 1.34), one of which was contrasted against a standard HIV prevention intervention.

META–REGRESSIONS

We used a stepwise elimination procedure to identify a core set of study characteristics associated with intervention effects (Table 3). After controlling for other characteristics, the most favorable effects were observed among older samples with
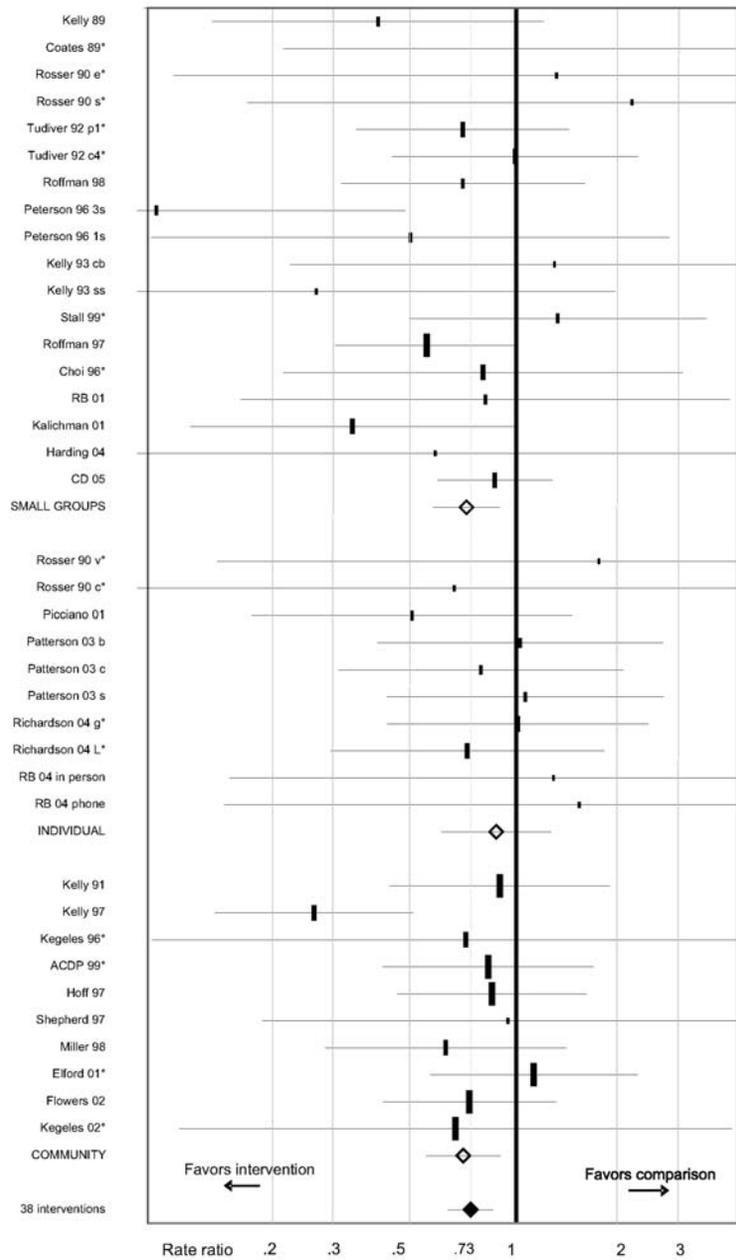
FIGURE 1. Effects of 38 HIV prevention interventions for men who have sex with men. Interventions compared to minimal or no HIV prevention control conditions, by intervention format and in approximate order of dates conducted. *Note.* Tick mark size is proportional to study weight. Suffixes after study year correspond to abbreviations used for intervention arms in Table 1. *Prevalence ratio substituted and confidence interval adjusted because rate ratio not available.
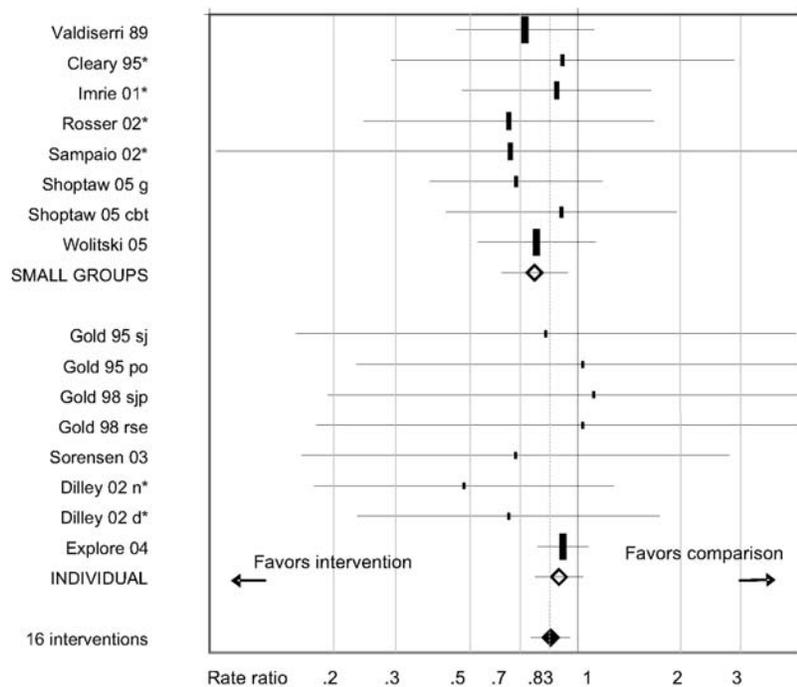
FIGURE 2. Effects of 16 HIV prevention interventions for men who have sex with men. Interventions compared with standard or other HIV prevention control conditions, by intervention format and in approximate order of dates conducted. *Note.* **Tick mark size is proportional to study weight. Suffixes after study year correspond to abbreviations used for intervention arms in Table 2. *Prevalence ratio substituted and confidence interval adjusted because rate ratio not available.**

more homogeneous ethnicity and lower prevalence of HIV. Statistically significant effects were obtained when participation in the assigned intervention was over 80% and in U.S. studies. Interventions measuring number of episodes of or partners for unprotected sex yielded somewhat more favorable results than those measuring only any unprotected sex versus none. The most favorable effects among small group interventions were those addressing perception of risk and losses ("unsafe sex exposes you") rather than gains ("safer sex protects you"). The most favorable effects among individual–level interventions were those that addressed losses and, among community–level interventions, those that addressed personal skills such as self–reinforcement for behavior change efforts.

EXPOSURE RATES IN COMMUNITY–LEVEL INTERVENTIONS

Exposure rates measure the proportion of the population that actually report receiving the intervention. High exposure rates were critical to the success of community–level interventions (Figure 3). Point estimates of effectiveness improved consistently from a prevalence ratio of 1.12 (favoring the comparison group) when

TABLE 3. Effects of 54 HIV Prevention Interventions with Men Who Have Sex With Men,
1988–2005 (stratified by study characteristics with mutual adjustment)

Variable/Level	Rate Ratio (95% CI)
Mean age	
21–29	.84 (.55, 1.18)
31–34	.74 (.58, .91)
35–42	.74 (.57, .92)
African American, Asian, Latino, Black, and Mulatto participants	
0–26%	.74 (.59, .88)
27–39%	.93 (.65, 1.23)
49–100%	.68 (.47, .91)
HIV prevalence	
0–11%	.71 (.53, .90)
14–61%	.76 (.60, .93)
100%	.77 (.54, 1.03)
Participation	
<65%	.80 (.57, 1.04)
74–76%	.93 (.65, 1.24)
82–100%	.74 (.56, .93)
Location	
United States	.71 (.57, .84)
Elsewhere	.87 (.62, 1.14)
Outcome measure	
Episodes of or partners for unprotected sex	.71 (.57, .85)
Any unprotected sex versus none	.81 (.63, .99)
Intervention format and content	
Small group	.72 (.57, .87)
Losses	.47 (.29, .67)
Perceived Risk	.57 (.41, .74)
Individual–level	.92 (.69, 1.16)
Losses	.76 (.40, 1.25)
Community–level	.59 (.38, .85)
Personal skills	.48 (.26, .77)

only 3% reported exposure to a prevalence ratio of 0.71 when 82% reported exposure. Because of the wide confidence intervals, the trend was not statistically significant ($p = .51$).

SENSITIVITY ANALYSES

Sensitivity analyses are performed to examine whether changes in assumptions have a major influence on results. The results reported above are robust to various changes in assumptions. If any one intervention among those compared against minimal to no HIV prevention controls had not been included, the result closest to null that would have been obtained among the remaining 37 interventions is a rate ratio of .78 (CI = .66, .91). Even if the seven interventions with the most favorable rate ratios were excluded, the summary effect of the remaining studies would still be favorable (RR = .84) and statistically significant (CI = .71, .99).

In the analyses presented above we assumed that the rate ratio equals the prevalence ratio when one or the other was not available. However regression models for the interventions for which both LnRR and LnPR were available suggest that LnRR may actually be 2.1 times the magnitude of LnPR. Sensitivity analysis adjusting for this factor yielded a rate ratio of .71 (CI = .60, .82) for the interventions tested

against minimal to no HIV prevention intervention. For the interventions tested against standard or other HIV prevention intervention, the adjusted rate ratio was .80 (CI = .70, .91). Thus the unadjusted rate ratios presented above may be slightly conservative.

Similarly, the typical LnPR may actually be only .37 times the magnitude of LnRR. Sensitivity analysis of prevalence ratios adjusting for this factor yielded an effect of .88 (CI = .82, .95) for the 32 interventions tested against minimal to no HIV prevention intervention, so results above concerning prevalence ratios for this group may be slightly overstated. Prevalence ratios were available for all 16 interventions tested against standard or other HIV prevention interventions so the adjustment factor has no impact on that subset.

Estimation of weights for rate ratios given only dichotomous data required an assumed value for the dispersion parameter d (see Appendix). For the analyses above, we used the geometric weighted mean value of 6.5 observed among the 19 studies for which this parameter could be estimated. The summary rate ratio for interventions compared with minimal to no HIV prevention controls became slightly more favorable (lower) with increasing assumed values of this parameter from .75 (CI = .65, .87) at a low value of $d = 3.3$, to .71 (CI = .60, .83) at a high value of $d = 15.7$. Because count data were available for most interventions that were compared with standard or other HIV prevention interventions, the summary rate ratio was essentially unaffected by varying values of d. Estimates of weights for prevalence ratios given summary statistics for count data do not involve an assumed value for d.

PUBLICATION BIAS

Meta–analysis may be vulnerable to publication bias if studies with less favorable results are not found and included. A useful test for publication bias is based on the funnel plot, which compares intervention effects with sample sizes (Macaskill et al., 2001). The typical thumbprint of publication bias is more favorable effects among small studies than among large studies. Modeled effects of interventions tested against minimal to no HIV prevention comparison conditions (controlling for intervention format and outcome metric [RR or PR]) were somewhat more favorable at the minimum sample size of 28 subjects (RR = 0.68) than at the maximum of 2324 subjects (RR = 0.79), but the trend was not significant (p for slope = .68). Similarly, modeled effects of interventions tested against standard or other HIV prevention interventions were more favorable at the minimum of 45 subjects (RR = 0.76) than at the maximum of 3775 (RR = 0.90, p for slope = .33). If the increasing slope is due to publication bias, our estimates may tend to overstate effects.

DISCUSSION

In studies with strong research designs, behavioral interventions for MSM reduced unprotected sex by 27% compared with minimal or no intervention and reduced the proportion of men reporting any unprotected sex by 16%. These statistically significant effects were also evident in subgroup analyses of small-group and community–level interventions.

However, it should not be assumed that simply doing anything is always better than nothing. Point estimates for these intervention effects are members of a distribution whose center indicates favorable effects but which includes some null and even a few slightly unfavorable results. As would be expected, less favorable effects were found among experimental interventions that were contrasted against standard or
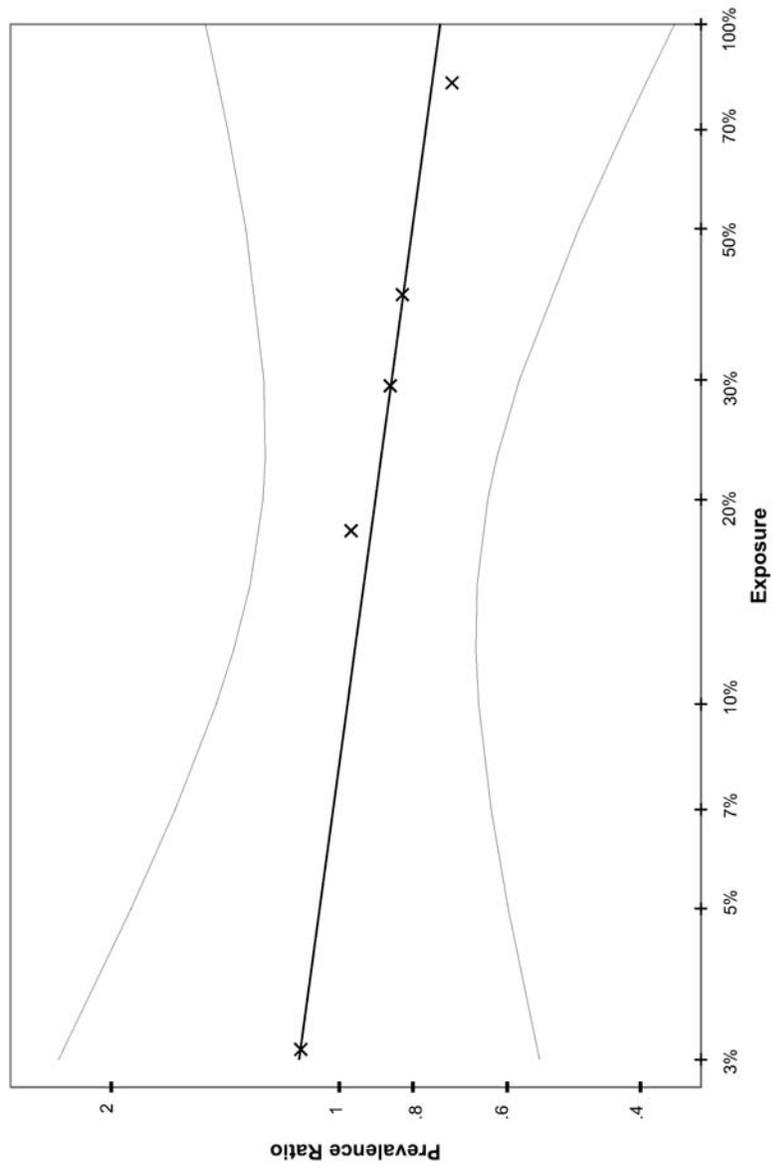
FIGURE 3. Intervention effects (prevalence ratios) by percentage reporting exposure in five community-level studies for men who have sex with men, with weighted linear regression and 95% confidence interval.

other HIV prevention interventions and among studies where the intended intervention was not effectively delivered to a substantial proportion (e.g., 30% or more) of the study sample.

The limited information on biological outcomes suggests that the highest risk clients may be better served by individual–level interventions than by small-group interventions that introduce them to potential new partners who are themselves at particularly high risk. In one study, a small group intervention for STD clinic patients showed modest reductions in unprotected sex but also resulted in more STD infections than a standard one–on–one counseling session about sexual risk behavior (Imrie et al., 2001). However an enhanced individual–level intervention with a similarly modest reduction in unprotected sex was accompanied by a substantial reduction in new HIV infections (EXPLORE Study Team, 2004). A third study of a small group intervention for HIV–positive MSM found moderate effect on behaviors but no effect on STDs (Wolitski et al., 2005).

Meta–analysis can be an essential tool for guiding future research. In terms of design, we found that count outcomes such as number of episodes of or partners for unprotected sex may be more sensitive than dichotomous outcomes, which do not recognize even a very large decrease in an individual's risk unless unprotected sex is altogether eliminated. A reduction in number of occasions of unprotected sex may have an important impact on HIV transmission, particularly if the number of partners for unprotected sex and the density of unprotected sexual networks also decrease.

Empirical examination of the effects of serosorting, negotiated safety, withdrawal before ejaculation, strategic positioning, and partner selection is urgently needed (Hoff, Faigeles, Wolitski, Purcell, Gomez, & Parsons, 2004). Because perceptions of partners' risk may not always be correct, the effectiveness of such strategies in avoiding HIV transmission is unknown. Availability of new treatments may contribute to complacency about HIV prevention (Demmer, 2003). In recent years the Internet has become an important factor in the HIV epidemic (Anonymous, 2004), but it may also be useful in prevention (Anonymous, 2003) and in partner notification (CDC, 2003). These relatively new factors that may influence behaviors and biological risk should be considered in future research.

Our review shows that some populations at high risk have been critically underserved in intervention research, particularly African American (Leone et al., 2003) and Latino MSM, and MSM in countries where English is not the primary language. Factors affecting HIV risk are likely to differ among such populations (Coleman, 2003; Courtenay–Quirk, Wolitski, Hoff, & Parsons, 2003; Millett, 2004; Zea, Reisen, & Diaz, 2003). Use of alcohol and drugs, particularly methamphetamine, and attending bathhouses, sex clubs, and circuit parties may be associated with risky sex among MSM (Crosby, DiClemente, & Mettey, 2003; Lister et al., 2003; Semple, Patterson, & Grant, 2003). A wide range of effective interventions is needed for those at highest risk.

APPENDIX: ESTIMATING VARIANCES OF INTERVENTION EFFECTS

RATE RATIOS

When summary statistics (sample sizes m and n, sample means $\hat{E}[Y]$ and $\hat{E}[X]$, and sample variances $V\hat{a}r[Y]$ and $V\hat{a}r[X]$) from count data are available, the variance of the natural logarithm of the rate ratio (LnRR) can be estimated by the delta method as:

$$V\hat{a}r\left(Ln\hat{R}R\right) = \frac{V\hat{a}r[Y]}{m\hat{E}[Y]^2} + \frac{V\hat{a}r[X]}{n\hat{E}[X]^2}$$

When only dichotomous outcome data (e.g., sample sizes m and n and sample prevalences $\hat{p}$ and $\hat{q}$) were available we substituted the value of the prevalence ratio for the rate ratio. In this case we used the method of moments to estimate the variance of LnRR as:

$$V\hat{a}r\left(Ln\hat{R}R\right) = \frac{\hat{d}(1-\hat{p})^{\hat{d}}}{m\left[1-(1-\hat{p})^{\hat{d}}\right]} + \frac{\hat{d}(1-\hat{q})^{\hat{d}}}{n\left[1-(1-\hat{q})^{\hat{d}}\right]}$$

For the pooled dispersion parameter $\hat{d}$ for the intervention and comparison groups we used the value 6.5, the weighted mean value of the dispersion parameters from 19 studies from which it could be directly estimated. The dispersion parameter is estimated as:

$$\hat{d} = \frac{V\hat{a}r[Z] - \hat{E}[Z]}{\hat{E}^2[Z]}$$

where $\hat{E}[Z]$ and $V\hat{a}r[Z]$ are the pooled sample mean and the pooled sample variance for all intervention and comparison groups in a study. This formula applies when the underlying count data are overdispersed relative to the Poisson distribution, that is, the variances are greater than the means.

PREVALENCE RATIOS

When summary statistics (sample sizes m and n and sample prevalences $\hat{p}$ and $\hat{q}$) are available from dichotomous data, the variance of the natural logarithm of the prevalence ratio (LnPR) can be estimated by the delta method as:

$$V\hat{a}r\left(Ln\hat{P}R\right) = \frac{1-\hat{p}}{m\hat{p}} + \frac{1-\hat{q}}{n\hat{q}}$$

When only count data (e.g., sample sizes m and n, sample means $\hat{E}[Y]$ and $\hat{E}[X]$, and sample variances $V\hat{a}r[Y]$ and $V\hat{a}r[X]$) were available we substituted the value of the rate ratio for the prevalence ratio. In this case we used the method of moments to estimate the variance of LnPR as:

$$V\hat{a}r\left(Ln\hat{P}R\right) = \frac{\hat{f}^j}{m\left(1-\hat{f}^j\right)} + \frac{\hat{g}^k}{n\left(1-\hat{g}^k\right)}$$

where

$$\hat{f} = \frac{\hat{E}(Y)}{V\hat{a}r(Y)}$$

$$\hat{j} = \frac{\hat{E}^2(Y)}{V\hat{a}r(Y) - \hat{E}(Y)}$$

$$\hat{g} = \frac{\hat{E}(X)}{V\hat{a}r(X)}$$

$$\hat{k} = \frac{\hat{E}^2(X)}{V\hat{a}r(X) - \hat{E}(X)}$$

$\hat{E}(Y) =$ intervention sample mean

$V\hat{a}r(Y) =$ intervention sample variance

$\hat{E}(X) =$ comparison sample mean

$V\hat{a}r(X) =$ comparison sample variance

This formula applies when the data are overdispersed, that is, the sample variances are greater than the sample means. Derivations and further justification of these formulas are available from the authors upon request.

REFERENCES

Anonymous. (2003). Interventions can ease dangers of cybersex. Counselors educate while in chatrooms. *AIDS Alert, 18,* 124–127.

Anonymous. (2004). Study confirms the link between MSM and web. San Francisco officials use on–line interventions. *AIDS Alert, 19,* 19–20.

Carballo–Diéguez, A., Dolezal, C., Leu, C.–S., Nieves, L., Díaz, F., Decena, C., et al. (2005). A randomized controlled trial to test an HIV–prevention intervention for Latino gay and bisexual men: Lessons learned. *AIDS Care, 17,* 314–328.

Catania, J.A., Morin, S.F., Canchola, J., Pollack, L., Chang, J., & Coates, T.J. (2000). U.S. priorities—HIV prevention. *Science, 290,* 717.

CDC AIDS Community Demonstration Projects Research Group. (1999). Community–level HIV intervention in 5 cities: Final outcome data from the CDC AIDS Community Demonstration Projects. *American Journal of Public Health, 89,* 336–345.

Centers for Disease Control and Prevention (2003). *HIV/AIDS Surveillance Report, 15.*

Centers for Disease Control and Prevention. (2004). Using the Internet for partner notification of sexually transmitted diseases—Los Angeles County, California, 2003. *Morbidity and Mortality Weekly Report, 53,* 129–131.

Chesney, M.A., Chambers, D.B., Taylor, J.M., Johnson, L.M., & Folkman, S. (2003). Coping effectiveness training for men living with HIV: Results from a randomized clinical trial testing a group–based intervention. *Psychosomatic Medicine, 65,* 1038–1046.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta–analysis. *Statistics in Medicine, 19,* 3127–331.

Choi, K.H., Lew, S., Vittinghoff, E., Catania, J.A., Barrett, D.C., & Coates, T.J. (1996). The efficacy of brief group counseling in HIV risk reduction among homosexual Asian and Pacific Islander men. *AIDS, 10,* 81–87.

Cleary, P.D., Van Devanter, N., Steilen, M., Stuart, A., Shipton–Levy, R., McMullen, W., et al. (1995). A randomized trial of an education and support program for HIV–infected individuals. *AIDS, 9,* 1271–1278.

Coates, T.J., McKusick, L., Kuno, R., & Stites, D.P. (1989). Stress reduction training changed number of sexual partners but not immune function in men with HIV. *American Journal of Public Health, 79,* 885–887.

Coleman, C.L. (2003). Determinants of HIV and AIDS among young African American men who have sex with men: A public health perspective. *Journal of National Black Nurses Association, 14,* 25–29.

Cooper, H., & Hedges, L.V. (1994). *The handbook of research synthesis.* New York: Russell Sage Foundation.

Conant, M. (2004). Losing the war on AIDS. *Journal of the American Academy of Dermatology, 51(1 Suppl),* S47–S48.

Courtenay–Quirk, C., Wolitski, R.J., Hoff, C., & Parsons, J.T. (2003). Interests in HIV prevention topics of HIV–seropositive men who have sex with men. *AIDS Education and Prevention, 15,* 401–412.

Crosby, R., DiClemente, R.J., & Mettey, A. (2003). Correlates of recent unprotected anal sex among men having sex with men attending a large sex resort in the South. *Sexually Transmitted Diseases, 30,* 909–913.

Deeks, J., on Behalf of the Statistical Methods Working Group of the Cochrane Collaboration (1999). *Statistical methods programmed in MetaView.* Oxford, England: Cochrane Library.

Demmer, C. (2003). HIV prevention in the era of new treatments. *Health Promotion Practice, 4,* 449–456.

Dilley, J.W., Woods, W.J., Sabatino, J., Lihatsh, T., Adler, B., Casey, S., et al (2002). Changing sexual behavior among gay male repeat testers for HIV: A randomized, controlled trial of a single–session intervention. *Journal of Acquired Immune Deficiency Syndrome, 30,* 177–186.

Elford, J., Bolding, G., & Sherr, L. (2001). Peer education has no significant impact on HIV risk behaviours among gay men in London. *AIDS, 15,* 535–538.

EXPLORE Study Team. (2004). Effects of a behavioural intervention to reduce acquisition of HIV infection among men who have sex with men: The EXPLORE randomised controlled study. *The Lancet, 364,* 41–50.

Flowers, P., Hart, G.J., Williamson, L.M., Frankis, J.S., & Der, G.J. (2002). Does bar–based, peer–led sexual health promotion have a community–level effect amongst gay men in Scotland? *International Journal of STD & AIDS, 13,* 102–108.

Garber, D.A., Silvestri, G., & Feinberg, M.B. (2004). Prospects for an AIDS vaccine: Three big questions, no easy answers. *Lancet Infectious Diseases, 4,* 397–413.

Gold, R.S., & Rosenthal, D.A. (1995). Preventing unprotected anal intercourse in gay men: A comparison of two intervention techniques. *International Journal of STD and AIDS, 6,* 89–94.

Gold, R.S., & Rosenthal, D.A. (1998). Examining self–justifications for unsafe sex as a technique of AIDS education: The importance of personal relevance. *International Journal of STD and AIDS, 9,* 208–213.

Gray, R.H., Li, X., Wawer, M.J., Gange, S.J., Serwadda, D., Sewankambo, N.K., et al (2003). Stochastic simulation of the impact of antiretroviral therapy and HIV vaccines on HIV transmission; Rakai, Uganda. *AIDS, 17,* 1941–1951.

Greenland, S. (1998). Meta–analysis. In K.J. Rothman, & S. Greenland (eds.), *Modern epidemiology* (pp. 643-673). Philadelphia: Lippincott Williams & Wilkins.

Harding, R., Bensley, J., Corrigan, N., Franks, L., Stratman, J., Waller, Z., & Warner, J. (2004). Outcomes and lessons from a pilot RCT of a community–based HIV prevention multi–session group intervention for gay men. *AIDS Care, 16,* 581–585.

Hasselblad, V., & Hedges, L.V. (1995). Meta–analysis of screening and diagnostic tests. *Psychological Bulletin, 117,* 167–78.

Hedges, L.V. (1994). Fixed effects models. In H. Cooper, & L.V. Hedges (eds.), *The handbook of research synthesis.* (pp. 285–299). New York: Russell Sage Foundation.

Hedges, L.V., & Vevea, J.L. (1998). Fixed– and random–effects models in meta–analysis. *Psychological Methods, 3,* 486–504.

Herbst, J.H., Sherba, R.T., Crepaz, N., DeLuca, J.B., Zohrabyan, L., Stall, R.D., et al. (2005). A meta–analytic review of HIV behavioral interventions for reducing sexual risk behavior of men who have sex with men. *Journal of Acquired Immune Deficiency Syndromes, 39,* 228–241.

Hoff, C.C., Faigeles, B., Wolitski, R.J., Purcell, D.W., Gomez, C., & Parsons, J.T. (2004). Sexual risk of HIV transmission is missed by traditional methods of data collection. *AIDS, 18,* 340–342.

Hoff, C.C., Kegeles, S.M., Acree, M., Stall, R., Paul, J., Ekstrand, M., & Coates, T.J. (1997). Looking for men in all the wrong places . . . : HIV prevention small–group programs do not reach high risk gay men. *AIDS, 11,* 829–830.

Imrie, J., Stephenson, J.M., Cowan, F.M., Wanigaratne, S., Billington, A.J., Copas, A.J., et al. (2001). A cognitive behavioural intervention to reduce sexually transmitted infections among gay men: Randomised trial. *British Medical Journal, 322,* 1451–1456.

Johnson, W.D., Hedges, L.V., & Diaz, R.M. (2003). Interventions to modify sexual risk behaviors

for preventing HIV infection in men who have sex with men (Cochrane Review). *The Cochrane Library,* no. 3, Chichester, England: Wiley.

Johnson, W.D., Hedges, L.V., Ramirez, G., Semaan, S., Norman, L.R., Sogolow, E., et al. (2002). HIV prevention research for men who have sex with men: A systematic review and meta–analysis. *Journal of Acquired Immune Deficiency Syndrome, 30*(Suppl. 1), S118–S129.

Johnson, W.D., Semaan, S., Hedges, L.V., Ramirez, G., Mullen, P.D., & Sogolow, E. (2002). A protocol for the analytical aspects of a systematic review of HIV prevention research. *Journal of Acquired Immune Deficiency Syndrome, 30*(Suppl. 1), S62–S72.

Kalichman, S.C., Rompa, D., Cage, M., DiFonzo, K., Simpson, D., Austin, J., et al (2001). Effectiveness of an intervention to reduce HIV transmission risks in HIV–positive people. *American Journal of Preventive Medicine, 21,* 84–92.

Kegeles, S.M., Hays, R.B., & Coates, T.J. (1996). The Mpowerment Project: A community–level HIV prevention intervention for young gay men. *American Journal of Public Health, 86,* 1129–1136.

Kegeles, S.M., Rebchook, G.M., Hays, R.B., & Pollack, L.M. (2002, July). *Staving off increases in young gay/bisexual men's risk behavior in the HAART era* [Abstract MoPeD3549]. Paper presented at the 14th International Conference on AIDS. Barcelona.

Kelly, J.A., Murphy, D.A., Bahr, R., Kalichman, S.C., Morgan, M.G., Stevenson, Y., et al. (1993). Outcome of cognitive–behavioral and support group brief therapies for depressed, HIV–infected persons. *American Journal of Psychiatry, 150,* 1679–1686.

Kelly, J.A., Murphy, D.A., Sikkema, K.J., McAuliffe, T.L., Roffman, R.A., Solomon, L.J., et al. (1997). Randomised, controlled, community–level HIV–prevention intervention for sexual–risk behaviour among homosexual men in US cities. *The Lancet, 350,* 1500–1505.

Kelly, J.A., St. Lawrence, J.S., Diaz, Y.E., Stevenson, L.Y., Hauth, A.C., Brasfield, T.L., et al (1991). HIV risk behavior reduction following intervention with key opinion leaders of population: An experimental analysis. *American Journal of Public Health, 8,* 168–171.

Kelly, J.A., St. Lawrence, J.S., Hood, H.V., & Brasfield, T.L. (1989). Behavioral intervention to reduce AIDS risk activities. *Journal of Consulting and Clinical Psychology, 57,* 60–67.

Leone, P., Hightow, L., Foust, E., Owen–O'Dowd, J., Phillip, S., Gray, P., et al. (2004). HIV transmission among black college student and non–student men who have sex with men — North Carolina, 2003. *Morbidity Mortal Weekly Report, 53,* 731–734.

Lister, N.A., Smith, A., Tabrizi, S., Hayes, P., Medland, N.A., Garland, S., & Fairley, C.K. (2003). Screening for Neisseria gonorrhoeae and Chlamydia trachomatis in men who have sex with men at male–only saunas. *Sexually Transmitted Diseases, 30,* 886–889.

Macaskill, P., Walter, S.D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta–analysis. *Statistics in Medicine, 20,* 641–654.

Miller, R.L., Klotz, D., & Eckholdt, H.M. (1998). HIV prevention with male prostitutes and patrons of hustler bars: Replication of an HIV preventive intervention. *American Journal of Community Psychology, 26,* 97–131.

Millett, G. (2004). More effort needed with "down low" males. CDC expert discusses issue in Q&A. *AIDS Alert, 19,* 42–43.

Murray, D.M. (1998). *Design and analysis of group–randomized trials.* New York: Oxford University Press.

O'Leary, A., DiClemente, R.J., & Aral, S.O. (1997). Reflections on the design and reporting of STD/HIV behavioral intervention research. *AIDS Education and Prevention, 9*(Suppl. A), 1–14.

Patterson, T.L., Shaw, W.S., & Semple, S.J. (2003). Reducing the sexual risk behaviors of HIV+ individuals: Outcome of a randomized controlled trial. *Annals of Behavioral Medicine, 25,* 137–145.

Perry, S., Fishman, B., Jacobsberg, L., Young, J., & Frances, A. (1991). Effectiveness of psychoeducational interventions in reducing emotional distress after human immunodeficiency virus antibody testing. *Archives of General Psychiatry, 48,* 143–147.

Peterson, J.L., Coates, T.J., Catania, J., Hauck, W.W., Acree, M., Daigle, D., et al. (1996). Evaluation of an HIV risk reduction intervention among African American homosexual and bisexual men. *AIDS, 10,* 319–325.

Picciano, J.F., Roffman, R.A., Kalichman, S.C., Rutledge, S.E., & Berghuis, J.P. (2001). A telephone based brief intervention using motivational enhancement to facilitate HIV risk reduction among MSM: A pilot study. *AIDS & Behavior, 5,* 251–262.

Richardson, J.L., Milam, J., McCutchan, A., Stoyanoff, S., Bolan, R., Weiss, J., et al. (2004). Effect of brief safer–sex counseling by medical providers to HIV–1 seropositive

patients: a multi–clinic assessment. *AIDS, 18,* 1179–1186.

Roffman, R.A., Picciano, J.F., Ryan, R., Beadnell, B., Fisher, D., Downey, L., et al. (1997). HIV–prevention group counseling delivered by telephone: an efficacy trial with gay and bisexual men. *AIDS and Behavior, 1,* 137–154.

Roffman, R.A., Stephens, R.S., Curtin, L., Gordon, J.R., Craver, J.N., Stern, M., et al. (1998). Relapse prevention as an interventive model for HIV risk reduction in gay and bisexual men. *AIDS Education and Prevention, 10,* 1–18.

Rosser, B.R.S. (1990). Evaluation of the efficacy of AIDS education interventions for homosexually active men. *Health Education Research, 5,* 299–308.

Rosser, B.R.S., Bockting, W.O., Rugg, D.L., Robinson, B.B., Ross, M.W., Bauer, G.R., et al. (2002). A randomized controlled intervention trial of a sexual health approach to long–term HIV risk reduction in men who have sex with men. Effects of the intervention on unsafe sexual behavior. *AIDS Education and Prevention, 14*(Suppl. 3A), 59–71.

Rotheram–Borus, M.J., Lee, M.B., Murphy, D.A., Futterman, D., Duan, N., Birnbaum, J.M. et al. (2001). Efficacy of a preventive intervention for youths living with HIV. *American Journal of Public Health, 91,* 400–405.

Rotheram–Borus, M.J., Swendeman, D., Comulada, W.S., Weiss, R.E., Lee, M., & Lightfoot, M. (2004). Prevention for substance–using HIV–positive young people: Telephone and in–person delivery. *Journal of Acquired Immune Deficiency Syndrome, 37,* S68–S77.

Sampaio, M., Brites, C., Stall, R., Hudes, E.S., & Hearst, N. (2002). Reducing AIDS risk among men who have sex with men in Salvador, Brazil. *AIDS and Behavior, 6,* 173–181.

Semaan, S., Kay, L., Strouse, D., Sogolow, E., Mullen, P.D., Neumann, M.S., et al. (2002). A profile of U.S.–based trials of behavioral and social interventions for HIV risk reduction. *Journal of Acquired Immune Deficiency Syndrome, 30*(Suppl. 1), S30–S50.

Semple, S.J., Patterson, T.L., & Grant, I. (2003). Binge use of methamphetamine among HIV–positive men who have sex with men: Pilot data and HIV prevention implications.

AIDS Education & Prevention, 15, 133–147.

Shepherd, J., Weare, K., & Turner, G. (1997). Peer–led sexual health promotion with young gay and bisexual men: Results of The HAPEER Project. *Health Education, 97,* 204–212.

Shoptaw, S., Reback, C.J., Peck, J.A., Yang, X., Rotheram–Fuller, E., Larkins, S., et al. (2005). Behavioral treatment approaches for methamphetamine dependence and HIV–related sexual risk behaviors among urban gay and bisexual men. *Drug and Alcohol Dependence, 78,* 125–349.

Sorensen, J.L., Dilley, J., London, J., Okin, R.L., Delucchi, K.L., & Phibbs, C.S. (2003). Case management for substance abusers with HIV/AIDS: A randomized clinical trial. *American Journal of Drug and Alcohol Abuse, 29,* 133–150.

Stall, R.D., Paul, J.P., Barrett, D.C., Crosby, G.M., & Bein, E. (1999). An outcome evaluation to measure changes in sexual risk–taking among gay men undergoing substance use disorder treatment. *Journal of Studies on Alcohol, 60,* 837–845.

Tudiver, F., Myers, T., Kurtz, R.G., Orr, K., Rowe, C., Jackson, E., et al. (1992). The Talking Sex Project: Results of a randomized controlled trial of small–group AIDS education for 612 gay and bisexual men. *Evaluation and the Health Professions, 15,* 26–42.

Valdiserri, R.O., Lyter, D.W., Leviton, L.C., Callahan, C.M., Kingsley, L.A., & Rinaldo, C.R. (1989). AIDS prevention in homosexual and bisexual men: results of a randomized trial evaluating two risk reduction interventions. *AIDS, 3,* 21–26.

Wolitski, R.D., Gómez, C.A., & Parsons, J.T., & the SUMIT Study Group. (2005). Effects of a peer–led behavioral intervention to reduce HIV transmission and promote serostatus disclosure among HIV–seropositive gay and bisexual men. *AIDS, 19(Suppl. 1),* S99–S109.

Zea, M.C., Reisen, C.A., & Diaz, R.M. (2003). Methodological issues in research on sexual behavior with Latino gay and bisexual men. *American Journal of Community Psychology, 31,* 281–291.

Appendix D. Behavioral Interventions to Reduce Risk for Sexual Transmission of HIV Among Men Who Have Sex with Men

This published Cochrane review is reproduced here with permission from the Cochrane Collaboration. This is a Cochrane review [Johnson WD, Diaz RM, Flanders WD, Goodman M, Hill AN, Holtgrave D, Malow R, McClellan, WM. Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men. *Cochrane Database of Systematic Reviews* 2008, Issue 3. Art. No.: CD001230. DOI: 10.1002/14651858.CD001230.pub2]. The published review in the Cochrane Library will be maintained and updated regularly as circumstances require. The version of the review included in this electronic copy of the dissertation will not be updated.

Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men (Review)

Johnson WD, Diaz RM, Flanders WD, Goodman M, Hill AN, Holtgrave D, Malow R, McClellan WM

THE COCHRANE COLLABORATION®

WILEY
Publishers Since 1807

TABLE OF CONTENTS

Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Wayne D Johnson[1], Rafael M. Diaz[2], William D Flanders[3], Michael Goodman[3], Andrew N Hill[3], David Holtgrave[3], Robert Malow[4], William M McClellan[3]

[1]Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, Atlanta, USA. [2]Center for Community Research, San Francisco State University, San Francisco, USA. [3]Rollins School of Public Health, Emory University, Atlanta, USA. [4]Stempel School of Public Health & AIDS Prevention Program, Florida International University, Miami, USA

Contact address: Wayne D Johnson, Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, Mailstop E-37, 1600 Clifton Road NE, Atlanta, GA, 30333, USA. wdj0@cdc.gov. (Editorial group: Cochrane HIV/AIDS Group.)

A B S T R A C T

Background

Men who have sex with men (MSM) remain at great risk for HIV infection. Program planners and policy makers need descriptions of interventions and quantitative estimates of intervention effects to make informed decisions concerning prevention funding and research. The number of intervention strategies for MSM that have been examined with strong research designs has increased substantially in the past few years.

Objectives

1. To locate and describe outcome studies evaluating the effects of behavioral HIV prevention interventions for MSM.

2. To summarize the effectiveness of these interventions in reducing unprotected anal sex.

3. To identify study characteristics associated with effectiveness.

4. To identify gaps and indicate future research, policy, and practice needs.

Search strategy

We searched electronic databases, current journals, manuscripts submitted by researchers, bibliographies of relevant articles, conference proceedings, and other reviews for published and unpublished reports from 1988 through December 2007. We also asked researchers working in HIV prevention about new and ongoing studies.

Selection criteria

Studies were considered in scope if they examined the effects of behavioral interventions aimed at reducing risk for HIV or STD transmission among MSM. We reviewed studies in scope for criteria of outcome relevance (measurement of at least one of a list of behavioral or biologic outcomes, e.g., unprotected sex or incidence of HIV infections) and methodologic rigor (randomized controlled trials or certain strong quasi-experimental designs with comparison groups).

Data collection and analysis

We used fixed and random effects models to summarize rate ratios (RR) comparing intervention and control groups with respect to count outcomes (number of occasions of or partners for unprotected anal sex), and corresponding prevalence ratios (PR) for dichotomous outcomes (any unprotected anal sex vs. none). We used published formulas to convert effect sizes and their variances for count and dichotomous outcomes where necessary. We accounted for intraclass correlation (ICC) in community-level studies and adjusted for baseline conditions in all studies. We present separate results by intervention format (small group, individual, or community-level) and by type of intervention delivered to the comparison group (minimal or no HIV prevention in the comparison condition versus standard or other HIV prevention in the comparison condition). We examine rate ratios stratified according to characteristics of participants, design, implementation, and intervention content. For small group and individual-level interventions we used a stepwise selection process to identify a multivariable model of predictors of reduction in occasions of or partners for unprotected anal sex. We used funnel plots to examine publication bias, and Q (a chi-squared statistic with degrees of freedom = number of interventions minus 1) to test for heterogeneity.

Main results

We found 44 studies evaluating 58 interventions with 18,585 participants. Formats included 26 small group interventions, 21 individual-level interventions, and 11 community-level interventions. Sixteen of the 58 interventions focused on HIV-positives. The 40 interventions that were measured against minimal to no HIV prevention intervention reduced occasions of or partners for unprotected anal sex by 27% (95% confidence interval [CI] = 15% to 37%). The other 18 interventions reduced unprotected anal sex by 17% beyond changes observed in standard or other interventions (CI = 5% to 27%).

Intervention effects were statistically homogeneous, and no independent variable was statistically significantly associated with intervention effects at alpha=.05. However, a multivariable model selected by backward stepwise elimination identified four study characteristics associated with reduction in occasions of or partners for unprotected anal sex among small group and individual-level interventions at alpha=.10. The most favorable reductions in episodes of or partners for unprotected anal sex (33% to 35% decreases) were observed among studies with count outcomes, those with shorter intervention spans (<=1 month), those with better retention in the intervention condition than in the comparison condition, and those with minimal to no HIV prevention intervention delivered to the comparison condition.

Because there were only 11 community-level studies we did not search for a multivariable model for community-level interventions. In stratified analyses including only one variable at a time, the greatest reductions (40% to 54% decreases) in number of episodes of or partners for unprotected anal sex among community-level interventions were observed among studies where groups were assigned randomly rather than by convenience, studies with shorter recall periods and longer follow-up, studies with more than 25% non-gay identifying MSM, studies in which at least 90% of participants were white, and studies in which the intervention addressed development of personal skills.

Authors' conclusions

Behavioral interventions reduce self-reported unprotected anal sex among MSM. These results indicate that HIV prevention for this population can work and should be supported.

Results of previous studies provide a benchmark for expectations in new studies. Meta-analysis can inform future design and implementation in terms of sample size, target populations, settings, goals for process measures, and intervention content.

When effects differ by design variables, which are deliberately selected and planned, awareness of these characteristics may be beneficial to future designs. Researchers designing future small group and individual-level studies should keep in mind that to date, effects of the greatest magnitude have been observed in studies that used count outcomes and a shorter intervention span (up to 1 month).

Among small group and individual-level studies, effects were also greatest when the comparison condition included minimal to no HIV prevention content. Nevertheless, statistically significant favorable effects were also seen when the comparison condition included standard or other HIV prevention content. Researchers choosing the latter option for new studies should plan for larger sample sizes based on the smaller expected net intervention effect noted above.

When effects differ by implementation variables, which become evident as the study is conducted but are not usually selected or planned, caution may be advised so that future studies can reduce bias. Because intervention effects were somewhat stronger (though not statistically significantly so) in studies with a greater attrition in the comparison condition, differential retention may be a threat to validity. Extra effort should be given to retaining participants in comparison conditions.

Among community-level interventions, intervention effects were strongest among studies with random assignment of groups or communities. Therefore the inclusion of studies where assignment of groups or communities was by convenience did not exaggerate the summary effect. The greater effectiveness of interventions including more than 25% non-gay identifying MSM suggests that when they can be reached, these men may be more responsive than gay-identified men to risk reduction efforts. Non-gay identified MSM may have had less exposure to previous prevention messages, so their initial exposure may have a greater impact.

The greater effectiveness of interventions that include efforts to promote personal skills such as keeping condoms available and behavioral self-management indicates that such content merits strong consideration in development and delivery of new interventions for MSM. And the finding that interventions were most effective for majority white populations underscores the critical need for effective interventions for MSM of African and Latino descent.

Further research measuring the incidence of HIV and other STDs is needed. Because most studies were conducted among mostly white men in the US and Europe, more evaluations of interventions are needed for African American and Hispanic MSM as well as MSM in the developing world. More research is also needed to further clarify which behavioral strategies (e.g., reducing unprotected anal sex, having oral sex instead of anal sex, reducing number of partners, avoiding serodiscordant partners, strategic positioning, or reducing anal sex even with condom use) are most effective in reducing transmission among MSM, the messages most effective in promoting these behaviors, and the methods and settings in which these messages can be most effectively delivered.

PLAIN LANGUAGE SUMMARY

Behavioral interventions can reduce unprotected sex among men who have sex with men (MSM).

Interventions to reduce unprotected sex include individual counseling, social and behavioral support (such as peer education, assertiveness and relationship support, discussing attitudes and beliefs, videos). Small group and community interventions include group counseling or workshops, interventions in community areas, training community leaders, and community-building empowerment activities. The review found that these behavioral interventions can lead to significant risk reduction in MSM.

Continued research is needed to identify which behavioral strategies are most effective in reducing transmission, and which intervention components are most effective in influencing those behaviors. More research is also needed on the most effective strategies for non-white MSM in wealthy countries, as well as for MSM in developing countries.

BACKGROUND

Behavioral prevention remains central to the effort to reduce HIV transmission. Although antiretroviral therapy has tremendous life-saving potential, it is expensive, does not cure, and may have debilitating side effects for some people [Conant 2004]. Risk behaviors may increase if people believe that new treatments reduce subsequent transmission [Gray 2003] or if a vaccine becomes available [Crosby 2006]. Recent vaccine trials have yielded discouraging results [Cohen 2007, Markel 2005, Garber 2004].

MSM continue to make up the largest proportion of new AIDS cases and HIV infections each year in Pattern I countries [UNAIDS.org; Catania 2000; Mills 1997]. Of the estimated 322,125 male adults and adolescents living with AIDS in the United States in 2005, 67% had been exposed through male-to-male sexual contact, including 8% who had been exposed through both male-to-male sexual contact and injection drug use [CDC 2007 page 8].

MSM are at high risk among all races and ethnicities. Among 30,956 cases of AIDS reported among men and male adolescents in the United States in 2005 [CDC 2007 table 19], 53% were MSM or MSM-IDU. By race and ethnicity, 73% of AIDS diagnoses reported among white men, 40% among black men, 49% among Hispanic men, 58% among Asian men, and 63% among American Indian or Alaska/Hawaii native men were MSM or MSM-IDU. These cases and new infections are concentrated in a group believed to constitute only 2% to 10% of the adult male population [Binson 1995].

And in contrast to trends reported in the 1990s, the burden among MSM is now increasing faster than among other populations. The estimated yearly number of new diagnoses of HIV or AIDS among MSM and MSM/IDU in the 33 states with named reporting increased by 11% from 17,699 in 2001 to 19,620 in 2005, while decreasing by 20% among all other people in these 33 states [CDC 2007 table 1].

MSM are at high risk for HIV infection in the developing world as well. A systematic review in low- and middle-income countries found a weighted average HIV prevalence of 12.8% among MSM [Baral 2007]. Compared to other reproductive-age adults, the odds of HIV infection among MSM were 33 times greater in Latin America, 19 times greater in Asia, and 3.8 times greater in Africa, but only 1.3 times greater in Eastern Europe, where contaminated injections play a critical role in the epidemic.

Unprotected anal sex remains the greatest risk factor identified for HIV transmission. A case-control study in Australia found that the odds of becoming infected with HIV were 57 times as great among men who reported receptive anal sex to ejaculation with casual partners without a condom as among men who did not [Read 2007]. However risks from sex with main partners, insertive sex, and sex without ejaculation were not ruled out.

Partner selection based on perceived serostatus is being used as a strategy for risk reduction among MSM but carries some risk [Rietmeijer 2007]. Among 2788 MSM ages 23-29 in 6 US cities, 267 (9.6%) had HIV and were not aware of their status [MacKellar 2007]. The proportion who were infected but not aware was particularly high among African American MSM at 28%, compared to 8% among Hispanic MSM and 4% among white MSM. Since those most recently infected may be most infectious and least likely to know of their status, reliance on partners' awareness and disclosure of their own serostatus may be a risky strategy [Wawer 2005].

Previous reviews of HIV prevention efforts have examined the effects of behavioral interventions across multiple populations at risk. Fisher and Fisher [Fisher JD 1992, Fisher JD 2006a, Fisher JD 2006b] concluded that critical intervention components included not only information but also motivation and skills. Choi and Coates [Choi 1996@] noted the importance of skills training, as well as a lack of intervention research for MSM of color, young MSM, and non-gay-identifying MSM. Holtgrave et al. [Holtgrave 1995, Holtgrave 2007] cited the need for sufficient resources, intensity, and cultural competency, and a basis in behavioral and social science theory and previous research. Oakley et al. [Oakley 1995] identified a need for stronger research designs. Stephenson et al. [Stephenson 2000] reported that successful interventions were characterized by extensive formative research or high attendance rates.

One review not specific to MSM included a meta-analysis of 12 intervention studies: Kalichman et al. [Kalichman 1996] found that intervention effects diminished across studies as time from intervention to follow-up increased from 1 to 6 months. One qualitative review focusing specifically on men who have sex with men [Kegeles 1998] noted that community-based interventions have the capacity to reach people who would not participate in facility-based interventions, and who may be at higher risk than many who enroll in small group or individual interventions.

Effects of behavioral interventions for MSM have now been evaluated in numerous randomized trials and strong quasi-experimental studies. Quantitative synthesis can help to optimize the usefulness and interpretability of results across studies. Our first meta-analyses of HIV prevention interventions for MSM through 1997 found a 26% reduction in unprotected anal sex compared to neutral or standard conditions [Johnson 2002a, and previous Cochrane review]. Our update to those reviews found a 27% reduction in unprotected anal sex in 38 interventions compared to neutral conditions and a 17% reduction in 16 interventions compared to standard conditions [Johnson 2005].

Another meta-analysis of interventions for MSM found favorable effects among 17 studies that reported a basis in behavioral theory (odds ratio = .65) and no effect among 3 that did not (odds ratio = 1.03) [Herbst 2005]. A meta-analysis of 19 studies meeting narrower criteria (e.g., MSM age 20 and older not known to be HIV-

positive, as well as sample size, retention rates, and other criteria of design, implementation, and reporting) also found significant favorable effects [Herbst 2007]. This Cochrane Review further updates the list of studies meeting the criteria of Johnson 2005.

OBJECTIVES

In this review we examine and summarize the behavioral effects of rigorously evaluated interventions for MSM. To the extent that data permit, we estimate several parameters needed by program planners and policymakers [Holtgrave 2000; Bulterys 1997].

Our research questions were:

1. What behavioral interventions to reduce risk of HIV transmission among MSM have been tested in randomized trials or in rigorously controlled quasi-experimental studies?

2. What populations have been served or underserved in these studies?

3. What are the effects of MSM interventions contrasted against minimal or no intervention comparison conditions?

4. What are the effects of MSM interventions contrasted against standard or other HIV prevention intervention conditions?

5. What are the effects of small group, individual-, and community-level interventions for MSM?

6. What characteristics of small group and individual-level intervention studies are most closely associated with magnitude of effects in a multivariable model?

7. What characteristics of community-level intervention studies are most closely associated with magnitude of effects in single-variable models?

METHODS

Criteria for considering studies for this review

Types of studies

We reviewed studies for scope based on types of participants (MSM) and interventions (behavioral interventions to prevent HIV or STDs). We reviewed studies in scope for relevance based on inclusion of specified outcome measures (HIV or STD incidence or HIV risk behaviors), and for methodological rigor based on study design (randomized controlled trials and certain quasi-experimental designs).

Non-randomized studies were considered eligible only if they included independent comparison groups where assignment to treatment status was not based on need or volition, and separate baseline measurements were also taken, as in the Untreated Control

Group Design with Pretest and Posttest [Cook & Campbell, pp 103-118]. Examples of studies that were *not* eligible were those that compared:
- people who chose to participate in an intervention to those who did not,
- baseline and follow-up measures with no separate comparison condition,
- only follow-up measures without baseline measures when either individuals or groups were assigned to treatment condition by a non-random process.

In a change from our previous Cochrane review, we did not include studies with the recurrent institutional cycle design which features data collection at only one time point for the comparison group. We excluded a non-randomized study in which large community agencies were chosen for intervention and small agencies for the control condition. We did not exclude studies on the basis of chance differences between intervention and comparison groups in demographics and baseline distribution of the outcome variable. We used results with appropriate statistical controls for such characteristics where available.

Types of participants

MSM regardless of age, race/ethnicity, sexual orientation (gay / homosexual, bisexual, heterosexual), gender identity (including transsexuals), nationality, etc. We included only studies in which MSM constituted at least one-third of the study sample (e.g., HIV-seropositives) or were specifically targeted by the intervention. When other populations were included, we either obtained outcome data for the MSM subset or reduced the study weight to reflect only the proportion who were MSM.

Types of interventions

Behavioral or social interventions designed to promote sexual risk reduction and thereby to reduce transmission of HIV or other STDs. These interventions may be delivered to individuals, small groups, or communities.

We excluded interventions that focused not on sexual transmission but on cognitive or affective outcomes such as distress associated with HIV testing, or health and coping for seropositive men [Perry 1991, Chesney 2003]. We also excluded pharmaceutical interventions.

Types of outcome measures

We included only studies that measured intervention effects on behaviors understood to affect risk of HIV transmission (e.g., unprotected sex, condom use, number of partners) and biologic outcomes including incidence of infection by HIV or other STD.

We defined unprotected sex as anal intercourse without a condom. Data concerning other sexual and drug use behaviors were frequently not available. Because unprotected anal sex is the most epidemiologically pertinent behavior for MSM [O'Leary 1997] and was available for all studies, we restricted our analyses to this outcome.

Dichotomous measures reflect the proportion of respondents reporting any unprotected sex during the recall period. Count-level outcomes reflect the number of occasions of unprotected sex or the number of partners for unprotected sex during the recall period. Methods for managing multiple outcomes are described below under Calculation of effect sizes and Statistical analyses.

HIV and STD incidence were reported in only a few studies so we did not perform quantitative analyses of these outcomes.

Search methods for identification of studies

We systematically reviewed the HIV prevention literature to find studies measuring the effects of behavioral interventions for MSM [Johnson 2002a, Sogolow 2002]. Resources included online databases (Medline, PsycInfo, PubMed, AIDSLine, Web of Science, ERIC, EMBASE, Social Science Citation Index, Applied Social Sciences Index and Abstracts, Cochrane Library Controlled Clinical Trials Register, the National Research Register, and the Computer Retrieval of Information on Scientific Projects [CRISP] database), reviews and other studies in the HIV prevention literature, expert recommendation, hand searches of journals (AIDS, AIDS and Behavior, AIDS Care, AIDS Education and Prevention, American Journal of Public Health, International Journal of STD & AIDS, Journal of Acquired Immune Deficiency Syndromes, Journal of the Association of Nurses in AIDS Care, AIDS Patient Care and STDs), and manuscripts and unpublished reports submitted by researchers. We did not restrict searches by country or language. The references of the eligible articles were also searched, a process that was iterated until no new references were identified. We also reviewed the citations from prior systematic reviews and meta-analyses for possible references. We also sent requests for information to researchers funded by National Institutes of Health (NIH), and contacted experts and agencies who could provide relevant materials.

Keywords for electronic searches varied according to database. As an example, a search of the 1996-2007 Medline database in December 2007 for (AIDS/prevention & control [pc] or HIV infections/pc or sexually transmitted diseases/pc) yielded 17,085 citations [Figure 1]. A search for (homosexuality or bisexuality or gay.mp or bisexual.mp or men who have sex with men.mp or HIV seropositivity/psychology) yielded 7967 citations, and a search for (random allocation or intervention studies or program evaluation or random.mp or randomize.mp or randomized.mp or randomly.mp) yielded 347,740 citations. Most quasi-experimental studies included the terms "intervention studies" or "program evaluation." There were 121 citations that were included in all three searches. Review of these 121 led to identification of 22 trials that were eligible by the criteria described below. Other search methods as well as analogous searches of other databases led to identification of 44 total studies evaluating 58 experimental interventions.

Figure 1.   0 Medline search.
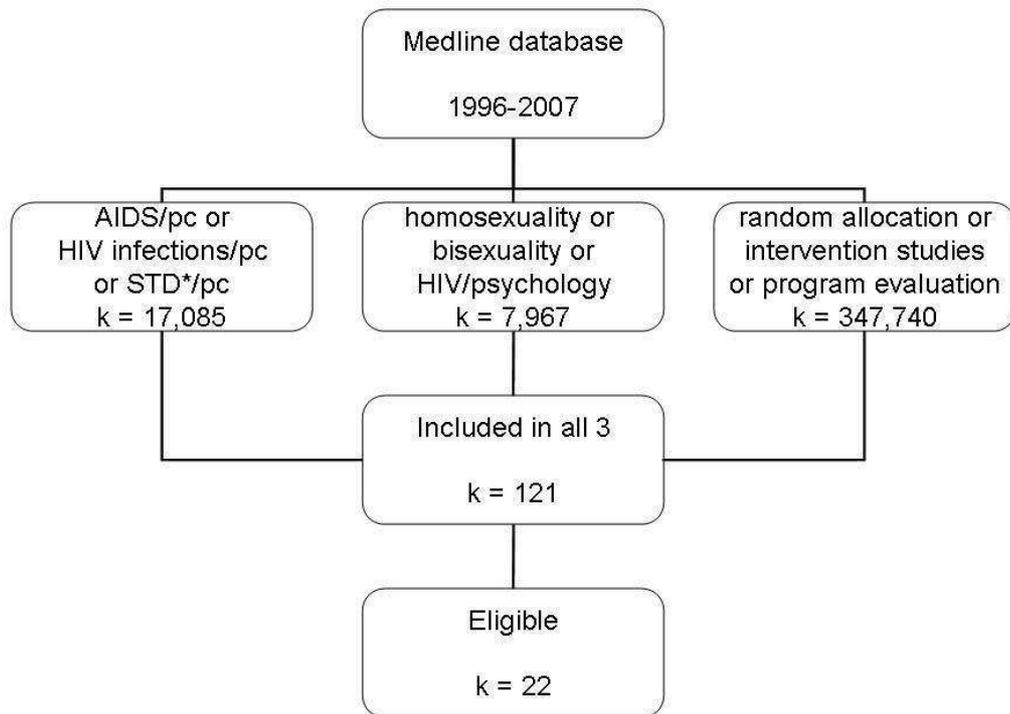
*Some keywords are abbreviated. Full keywords are provided in the text

Data collection and analysis

Studies found relevant and rigorous were eligible for the review. Data concerning outcomes, details of the interventions, and other study characteristics were independently abstracted from relevant studies by two reviewers using standardized data abstraction forms. Discrepancies were resolved by discussion and consensus.

Calculation of effect sizes

For each study, we calculated two effect sizes: a rate ratio (RR) and a prevalence ratio (PR). Because eligible studies used randomized or quasi-experimental designs, rate ratios (RR) can be used to estimate intervention effects for count measures and prevalence ratios (PR) for dichotomous measures [Cochrane Reviewers' Handbook; Deeks 2002]. We could have chosen standardized mean differences and odds ratios instead, but rate ratios and prevalence ratios have the advantage of being directly interpreted as 1 minus the net change. For example, a rate ratio of .73 indicates a 27% reduction in episodes of or partners for unprotected sex in the intervention condition after accounting for change in the comparison condition.

For each study that reported count measures, the rate ratio at follow-up was the ratio of the mean number of occasions of or partners for unprotected sex in the intervention group to the mean number in the comparison group. Similarly the prevalence ratio at follow-up was the ratio of the proportion of respondents reporting unprotected sex in the intervention group to the proportion in the comparison group. Rate ratios and prevalence ratios less than one represented a difference favoring the experimental intervention group. For each measure, the natural logarithm (LnRR or LnPR) was then an estimate of the intervention effect. The reciprocal of the variance of the logarithm of the measure served as a measure of the weight of information provided by the study.

When individual-level data were available, we used SAS Proc Genmod (except for group-randomized trials) to estimate rate ratios and prevalence ratios adjusted for the baseline value of covariates such as the outcome variable, age, race/ethnicity, and serostatus. For rate ratios, we used the negative binomial distribution and the log link function and adjusted the scale for Pearson's chi-square divided by its degrees of freedom. For prevalence ratios, we used the binomial distribution and the log link function.

When individual-level data or adjusted statistics were not available, we adjusted for the baseline distribution of the outcome variable by subtracting the baseline effect size (LnRR or LnPR) from the follow-up effect size. We used the lesser of the baseline and follow-up weights for such studies.

In studies where communities are the unit of assignment to treatment, the variance of the intervention effect will be underestimated and the weight (the reciprocal of the variance) will be overestimated if intraclass correlation (ICC) is not accounted for [Murray 1998]. We derived the adjustment factor to reduce study weights where necessary to account for ICC. For small values of ICC, the adjustment factor is approximately equal to Donner's variance inflation factor (VIF): $1 + ICC \times (m - 1)$ where m is the number of subjects in each unit of assignment. We assumed an ICC of .005, the value observed in the one study for which ICC was published [Kelly 1997].

We used outcome variables that did not distinguish between insertive and receptive sex, main and nonmain partners, or partners perceived to be seroconcordant vs. serodiscordant when such data were available. For studies from which the only available results were separated by insertive vs. receptive sex, or main vs. nonmain partners, we used the average point estimate and the average weight of the two measures to estimate the underlying combined effect [Johnson 2002b]. We accepted results concerning only nonmain, serodiscordant, or unknown serostatus partners when results were not available concerning main or seroconcordant partners.

For studies that compared two or more experimental interventions against a single control group, we allocated the control group into equal parts for comparison to each of the interventions. This strategy uses each individual's response only once and is thus valid for calculation of summary effects in fixed effects models. If this strategy had been necessary for several large studies, it could bias heterogeneity statistics toward a finding of homogeneity. If results had been heterogeneous, necessitating use of random effects models, the apparent variances could be understated. However the strategy was necessary for only a few small studies, and results would still have been homogeneous if these studies were left out altogether.

For studies that reported outcomes at multiple time points, the most commonly used follow-up times were 6 and 12 months after the end of the intervention. In order to focus on more sustained intervention effects, we selected outcomes measured closest to 12 months after intervention.

Statistical Analysis

We applied the standard procedures for meta-analysis to conduct summary, stratified, and regression analyses [Hedges 1994]. We conducted separate summary meta-analyses for rate ratios and prevalence ratios. In order to include all studies in each analysis, we substituted prevalence ratios for rate ratios in studies that measured only dichotomous outcomes (these are indicated by the at-sign @ in citations below, in the table of included studies, and in figures), and vice versa in studies that measured only count outcomes (these are indicated by the pound-sign #). If the prevalence ratio is constant across cutpoints then the rate ratio equals the prevalence ratio. This assumption of constancy across cutpoints is analogous to the assumption used to justify transformation between log odds ratios and standardized mean differences in other meta-analyses [Hasselblad 1995, Chinn 2000, Johnson 2002b].

Variances (and therefore weights) differ substantially between count and dichotomous outcomes and are not interchangeable between rate ratios and prevalence ratios. We used the method of moments to develop an estimate of the variance of LnRR when only dichotomous data were available, and the variance of LnPR when only count data were available [Johnson 2005].

Because intervention effects were highly homogeneous, random- and fixed-effects models yielded identical results for most analyses [Hedges 1998, DerSimonian 1986]. Variance estimates under the random effects model are always greater than or equal to those observed under the fixed effects model; in this sense the random effects model is conservative. But estimates of the summary intervention effect can be exaggerated by the random effects model in some circumstances, and in this sense the random effects model can be even more strongly anti-conservative. In our analyses, summary variances were almost indistinguishable between the two models, but summary effect estimates were occasionally farther from the null (smaller) under the random effects model. Thus the results according to fixed effects models are more conservative for these data, and they are the only ones we present.

We hypothesized *a priori* that interventions contrasted against neutral comparison conditions (minimal or no treatment related to HIV risk reduction) would yield stronger effects than those contrasted against active comparison conditions (standard or other HIV prevention interventions). Therefore we present results separately according to type of comparison condition.

Behavioral interventions for HIV prevention are frequently categorized by format as small group, individual-level, or community-level interventions. Therefore we present subgroup analyses for each of these categories. Interventions that included both a small group component and an individual-level component were classified as small group interventions. Community-level interventions that included small group or individual-level components were classified as community-level.

We examined differences in effectiveness for small group and individual-level interventions according to characteristics of participants, design, implementation, and intervention content, with statistical control for comparison type (minimal to no HIV prevention in the comparison condition versus standard or other HIV prevention in the comparison condition).

We used a backward stepwise selection process to identify a core set of study characteristics associated with intervention effects among small group and individual-level interventions. We excluded the "random assignment" variable from the selection process because the direction of the effect was counter to the *a priori* hypothesis that non-randomized trials might yield spuriously strong effects.

Because of the smaller number of community-level studies we did not conduct a multivariable analysis with this group. Because the list of variables associated with effectiveness was different for community-level studies we did not combine them in the multivariable analysis of small group and individual-level studies. For the sake of brevity, we present stratified and multivariable analyses only for rate ratios and not for prevalence ratios.

To investigate the possibility of publication bias, we examined funnel plots of treatment effect by sample size [Macaskill 2001].

R E S U L T S

Description of studies

See: Characteristics of included studies; Characteristics of excluded studies.

As of December 2007, we had identified 44 eligible studies evaluating 58 experimental HIV prevention interventions for MSM [Table of Included Studies]. Primary citations for these studies were found in 21 journals and 1 conference proceeding, with the largest numbers published in AIDS (9), the American Journal of Public Health (5), and the Journal of Acquired Immune Deficiency Syndromes (4). Two journals (the International Journal of STD and AIDS, and AIDS and Behavior) each published primary citations for 3 studies. Eleven studies tested two or more experimental interventions against comparison conditions, in which case we allocated the control group into equal parts for comparison to each of the interventions as described above. We treat the total of 58 experimental interventions and their paired control conditions as separate units for description and analyses.

Most of the interventions (k [number of interventions] = 40) were compared to minimal or no HIV prevention control conditions. Of these, 18 interventions were delivered in small group format, 11 in individual-level format, and 11 in community-level format. The other 18 interventions (which are marked with an asterisk in the tables and references), including 8 small group interventions and 10 individual-level interventions, were compared to standard or other HIV prevention conditions.

Over two-thirds of the 58 interventions (k = 41) were evaluated in the United States. Also represented were the United Kingdom (k = 5), Australia (k = 4), New Zealand (k = 4), Canada (k = 2), Brazil (k =1) and an international study in Russia and Bulgaria (k=1). The weighted mean age of participants was 33 years (range 21 to 42).

Across all studies, an average of 30% of participants were African American, Latino, Asian, or of other race/ethnic groups besides whites. Only 4 interventions in 3 studies focused on specific racial or ethnic groups: African Americans [Peterson 1996 1s; Peterson 1996 3s], Asians and Pacific Islanders [Choi 1996@], and Latinos [Carballo-Diéguez2005]. In 6 US studies evaluating 8 interventions [Cleary 1995*@; Kalichman 2001#; Rotheram-Borus 2001@; Richardson 2004 g@; Richardson 2004 L@; R-B 2004 in person@; R-B 2004 phone@; Healthy Living 2007#], the majority of participants were of race/ethnic groups other than white. In the study in Brazil [Sampaio 2002*@], 51% identified themselves as mulatto, 34% as white, and 15% as black. In 7 more US studies evaluating 10 interventions [Kelly 1993 cb; Kelly 1993 ss; Miller 1998; CDC ACDP 1999@; Patterson 2003 b#; Patterson 2003 c#; Patterson 2003 s#; Wolitski 2005*; Read 2006*#; Dilley 2007*#], more than a third of participants were ethnic minorities. Sixteen interventions focused on HIV-positive populations. In 4 of these interventions [Coates 1989@; Kelly 1993 cb; Kelly 1993 ss; Wolitski 2005*], all or nearly all (94% to 100%) partic-

ipants were MSM. In the other studies of HIV-positives [Cleary 1995*@; Kalichman 2001#; Rotheram-Borus 2001@; Sorensen 2003*; Patterson 2003 b#; Patterson 2003 c#; Patterson 2003 s#; Richardson 2004 g@; Richardson 2004 L@; R-B 2004 in person@; R-B 2004 phone@; Healthy Living 2007#], the largest subset (49% to 80%) of participants were MSM. HIV prevalence was particularly high in 3 other studies evaluating 4 interventions: Shoptaw 2005 cbt* and Shoptaw 2005 g* (61%), Stall 1999@ (50%), and Carballo-Diéguez2005 (36%).

Risk of bias in included studies

Assignment to treatment condition was random for most (48 of 58) interventions. For 7 interventions, large units such as cities, neighborhoods, or clinics were assigned to treatment condition based on convenience. Assignment was alternated between conditions in 3 other studies.

Among the small group and individual-level interventions, overall retention was high (from 80 to 100%) for 24 interventions and low (38 to 79%) for 21 interventions. Retention rates were not available for 2 studies; these were combined with the high retention rate group for analyses below. Also among the small group and individual-level interventions, 15 studies had greater retention in the comparison condition (mean = 8.5% greater, standard deviation = 7.6%), while 19 studies had greater retention in the intervention condition (mean = 8.4% greater, standard deviation = 9.0%). For analyses of differential retention as a dichotomous variable, 5 studies for which retention was equal in the two conditions were combined with the 15 studies with greater retention in the comparison condition, and 8 studies for which differential retention was not available were combined with the 19 studies with greater retention in the comparison condition.

Effects of interventions

EFFECTS OF INTERVENTIONS VS. MINIMAL TO NO HIV PREVENTION
The 40 interventions that were measured against minimal to no HIV prevention intervention reduced the number of episodes of or partners for unprotected sex by 27% (95% confidence interval [CI] = 15%, 37%) [Table 01.01]. The total MSM sample size in these 40 interventions was 11,864. The corresponding rate ratio was .73 (CI = .63, .85). This effect represents a decrease from the average background mean of 10.1 unprotected occasions in a 6-month period to 7.4 (CI = 6.4, 8.6), and from 1.2 partners for anal sex without condoms in a 6-month period to 0.9 (CI = 0.8, 1.0). The intervention effects were statistically homogeneous (Q[39 df] = 28.3, p for test of heterogeneity = 0.90).

In subgroup analyses the reduction was 30% (CI = 10%, 45%) among 18 small group interventions, 20% (CI = -6%, 40%) among 11 individual-level interventions, and 30% (CI = 9%,

45%) among 11 community-level interventions. Effects within each intervention format were also quite homogeneous (p = 0.74 for small groups, 0.99 for individual-level, and 0.29 for community-level interventions) indicating that results were statistically consistent among interventions within each format.
The same 40 interventions reduced the proportion of subjects reporting unprotected sex by 23% (CI = 17%, 28%) [Table 01.02]. The corresponding prevalence ratio was .77 (CI = .72, .83). This effect represents a decrease from an average of 41% reporting unprotected sex to 32% (CI = 30%, 34%). In subgroup analyses, significant reductions in the proportion reporting unprotected sex were observed for all three subgroups: a 27% reduction for small group interventions (CI = 16%, 36%), 16% for individual-level interventions (CI = 3%, 26%), and 25% for community-level interventions (CI = 16%, 34%). Again, the effects among studies were consistent both overall (p for heterogeneity = .14) and within subsets (p = .16 for small groups, p = .15 for individual-level, and p = .57 for community-level interventions).
EFFECTS OF INTERVENTIONS VS. STANDARD OR OTHER HIV PREVENTION
The 18 remaining interventions reduced the number of episodes of or partners for unprotected sex by 17% beyond changes observed in standard or other HIV prevention interventions (CI = 5%, 27%) [Table 02.01]. The total MSM sample size in these 18 studies was 6721. The corresponding rate ratio was .83 (CI = .73, .95). In subset analyses, the reductions were 23% (CI = -1%, 41%) among 8 small group interventions and 14% (CI = 0%, 27%) among 10 individual-level interventions. There were no community-level interventions in this subset.
The same 18 interventions reduced the proportion reporting unprotected sex by 7% beyond changes observed in standard or other HIV prevention interventions (CI = 3%, 11%) [Table 02.02]. The corresponding prevalence ratio was .93 (CI = .89, .97). The reduction was 13% among the 8 small group interventions (CI = 3%, 22%), and 6% among the 10 individual-level interventions (CI = 1%, 10%).
STRATIFIED ANALYSES OF RATE RATIOS FOR SMALL GROUP AND INDIVIDUAL-LEVEL INTERVENTIONS
Summary effects for all subgroups were in the favorable direction, but effects were statistically significant for some subgroups and not for others [Table 1]. We hypothesized *a priori* that intervention effects would be strongest in studies with a neutral comparison condition (minimal to no HIV prevention content). Therefore all analyses in this table (except for comparison condition itself) are controlled for type of comparison condition. For each set of stratification variables (participants, design, implementation, and intervention content), we note the stratum of studies that yields the most favorable results in terms of the point estimate of the percentage decrease in risky behavior (1 minus the rate ratio).

Table 1. Stratified analyses: small grp & individual intvn controlled for comparison type

Variable	Level	k	RR (95% CI)
OVERALL		47	.80 (.72, .89)*
PARTICIPANTS			
Location	US	34	.78 (.69, .88)*
	Elsewhere	13	.87 (.61, 1.25)
Mean age	21-33	21	.73 (.58, .93)*
	34-42	26	.80 (.71, .91)*
% Race/ethnic minority	4-24%	22	.75 (.61, .92)*
	35-100%	25	.81 (.70, .93)*
% HIV positive	0-22%	27	.77 (.66, .89)*
	36-100%	20	.82 (.67, 1.01)
% Non-gay identified	0-18%	27	.79 (.68, .91)*
	20-47%	20	.78 (.62, .98)*
DESIGN			
Comparison condition**	No HIV prevention	29	.74 (.62, .89)*
	Standard/other HIV prevention	18	.83 (.73, .95)*
Allocation	Random	44	.78 (.69, .88)*
	Not random	3	.95 (.51, 1.74)
Span	<1 month	25	.69 (.56, .85)*
	1.6-12 months	22	.83 (.72, .94)*
Peer delivery	No	40	.80 (.71, .90)*
	Yes	7	.71 (.52, .97)*

Table 1. Stratified analyses: small grp & individual intvn controlled for comparison type *(Continued)*

Duration	<=30 minutes	24	.74 (.59, .91)*
	>= 1 hour	23	.81 (.71, .92)*
Outcome measure	count	27	.77 (.68, .88)*
	dichot only	20	.84 (.65, 1.09)
Years conducted	1986-1996	24	.75 (.59, .97)*
	1997-2004	23	.80 (.70, .92)*
Group size	1 (individual)	21	.82 (.71, .96)*
	6-9	11	.76 (.59, .96)*
	10-25	15	.71 (.55, .93)*
Recall period	<3 months	17	.76 (.58. .99)*
	3-6 months	30	.79 (.70, .90)*
Time to followup	<4 months	21	.77, (.62, .95)*
	5-12 months	26	.80 (.69, .92)*
IMPLEMENTATION			
Retention better in	Intervention	27	.67 (.51, .88)*
	Comparison or equal	20	.81 (.72, .92)*
Overall retention	>=80%	26	.82 (.71, .94)*
	<80%	21	.73 (.60, .90)*
Participation	37-64%	20	.82 (.70, .96)*
	80-100%	27	.74 (.62, .89)*
Background prevalence	13-35%	23	.81 (.64, 1.02)
	37-73%	24	.78 (.68, .89)*

INTERVENTION CONTENT			
Losses		12	.62 (.44, .88)*
Self-esteem		15	.74 (.56, .96)*
Gains		18	.77 (.60, .99)*
Responsibility		23	.79 (.68, .92)*
Risk perception		22	.79 (.68, .93)*
Information		22	.80 (.64, .99)*
Technical skills		19	.80 (.59, 1.09)
Personal skills		32	.81 (.71, .92)*
Interpersonal skills		29	.82 (.72, .93)*
Self-efficacy		18	.83 (.71, .96)*
Social acceptability		16	.84 (.09, 8.07)*
Other content		7	.92 (.74, 1.13)
* p<.05			
** all other models controlled			
for comparison condition			

Studies performed in the United States yielded a 22% reduction in number of episodes of or partners for unprotected sex that was statistically significant (CI = 12%, 31%) while those performed elsewhere showed a net reduction of 13% that was not statistically significant (CI = -25%, 39%). Summary effects for all other subgroups according to participant characteristics were favorable and statistically significant. The most favorable summary effect by type of participants, a 27% reduction in risky behavior, was observed in studies with younger mean age (21 to 33 years).

All stratifications by characteristics of design resulted in statistically significant favorable results except for two groupings: studies in which assignment was not random, and studies for which a count outcome for unprotected sex was not available (in which case the dichotomous unprotected sex outcome was used instead with transformations as described above). The most favorable summary effect by design features, a 31% reduction in risky behavior, was observed in studies with a shorter intervention delivery time (up to one month).

With respect to characteristics of implementation, all subsets yielded statistically significant effects except for those with a lower background prevalence of unprotected sex. The most favorable effect by implementation features, a 33% reduction in risky behavior, was observed among those studies where retention of participants at followup was better in the intervention condition than in the comparison condition.

Finally, summary effects of interventions including each type of

content were statistically significant except for those including technical skills and those including "other" content. The most favorable effect by intervention content, a 38% reduction in risky behavior, was observed among interventions addressing perception of risk and losses ("unsafe sex puts you at risk") rather than gains ("safer sex protects you").

MULTIVARIABLE MODEL OF RATE RATIOS FOR SMALL GROUP AND INDIVIDUAL-LEVEL INTERVENTIONS

A multivariable model of rate ratios for small group and individual-level interventions was selected by a backwards elimination process [Table 2]. All variables described in Table 1 were included in the first model, and the variable with the smallest effect was removed. This process was repeated until each of the variables was retained with p<.10. The four variables remaining in the model all pertained to design and implementation. The most favorable reductions in episodes of or partners for unprotected sex among small group and individual-level interventions (33% to 35%) were observed among studies with count outcomes, shorter intervention span (<= 1 month), better retention in the intervention group than in the comparison group, and minimal to no HIV prevention delivered to the comparison condition.

Table 2. Multivariable model: Small group and individual-level interventions

Variable	Level	k	RR (95% CI)
Outcome type	Count	27	.65 (.53, .79)*
	Dichot only	20	.84 (.64, 1.10)
Span of intervention	<=1 month	25	.65 (.53, .82)*
	>1 month	22	.84 (.69, 1.02)
Retention better in	Intervention	27	.65 (.50, .86)*
	Comparison or about equal	20	.83, (.71, .98)*
Control condition	No HIV prevention	29	.67 (.54, .84)*
	Some HIV prevention	18	.81 (.67, .98)*
*p<.05			

By contrast the least favorable reductions (16% to 19% decreases) were observed among studies with only dichotomous outcomes, those with longer intervention spans, those with approximately equal retention or better retention in the comparison condition, and those where the comparison condition received some intervention relevant to HIV prevention. However the summary effects were still significant even among studies with stronger design and implementation characteristics, specifically, HIV prevention content delivered to the comparison condition and less attrition from the comparison condition.

STRATIFIED ANALYSES OF RATE RATIOS FOR COMMUNITY-LEVEL INTERVENTIONS

We also examined differences in effectiveness for community-level interventions according to characteristics of participants, design, implementation, and intervention content [Table 3]. The comparison type for all community-level interventions was minimal to no HIV prevention, so statistical control for comparison type was not necessary in this subgroup.

Table 3. Stratified analyses for community-level interventions

Variable	Level	k	RR (95% CI)
OVERALL		11	.70 (.55, .91)*
PARTICIPANTS			
% Non-gay identified	0-21%	6	.86 (.61, 1.23)
	26-100%	5	.59 (.42, .82)*
% Race/Ethnic minority	0-10%	5	.60 (.42, .87)*
	14-39%	6	.81 (.57, 1.15)
Location	US	7	.63 (.47, .86)*
	UK, Russia, Bulgaria	4	.87 (.56, 1.34)
Mean age	22-30	5	.85 (.46, 1.55)
	31-35	6	.68 (.51, .89)*
% HIV positive	1-6%	5	.80 (.45, 1.45)

Table 3. Stratified analyses for community-level interventions (*Continued*)

	7-16%	6	.68 (.52, .90)*
DESIGN			
Allocation	Random	4	.46 (.23, .90)*
	Not random	7	.83 (.62, 1.11)
Recall period	1-2 months	6	.56 (.40, .81)*
	3-12 months	5	.88 (.61, 1.25)
Follow-up time	<=2 months	5	.82 (.58, 1.16)
	6-16 months	6	.59 (.40, .85)*
Span (months)	1-6	6	.62 (.43, .88)*
	8-32	5	.80 (.56, 1.14)
Outcome type	Count	6	.63 (.47, .85)*
	Dichotomous	5	.92 (.57, 1.48)
Date conducted	1989-1994	5	.64 (.45, .89)*
	1996-2004	6	.80 (.55, 1.16)
Sampling	Serial cross-section	6	.66 (.50, .88)*
	Cohort	5	.86 (.51, 1.47)
IMPLEMENTATION			
Background prevalence of un-protected sex	11-29%	5	.61 (.41, .90)*
	30-68%	6	.78 (.56, 1.09)
Exposure to intervention	3-30%	3	.88 (.61, 1.27)
	40-82%	3	.79 (.42, 1.51)
INTERVENTION CONTENT			

Table 3. Stratified analyses for community-level interventions *(Continued)*

Personal skills		3	.60 (.40, .91)*
Gains		9	.66 (.50, .88)*
Risk perception		6	.67 (.49, .91)*
Interpersonal skills		8	.68 (.50, .92)*
Information		7	.82 (.58, 1.14)
Technical skills		4	.85 (.54, 1.33)
p<.05			

Summary effects for all subgroups of community-level interventions were favorable, but effects were statistically significant for some subgroups and not for others. With regard to characteristics of participants, the greatest reductions in unprotected sex (reductions of 40% or more) were observed among studies with more than 25% non-gay identified men and those whose samples were at least 90% white. Significant reductions were also observed among studies conducted in the US; those with mean age of 31 or older; and those with 7% or more HIV-positive men.

In terms of design and implementation, significant reductions of 40% or more were observed among community-level interventions with random allocation, shorter recall periods, and longer followup times. Significant summary effects were also seen among studies with shorter intervention spans, count outcomes, those conducted before 1995, those with serial cross-sectional designs, and those with a lower background prevalence of unprotected sex showed significant summary results. Summary effects in other strata of design and implementation were not significant.

In terms of intervention content, a significant 40% reduction in unprotected sex was observed among studies that addressed personal skills. Significant reductions were also seen among studies that addressed gains, risk perception, and interpersonal skills. Fa-

vorable but non-significant reductions were observed among studies that addressed information and technical skills.

Because there were only 11 community-level interventions studies, we did not search for a multivariable model for this subset.

PUBLICATION BIAS

Meta-analysis may be vulnerable to publication bias if studies with less favorable results are not found and included. A useful test for publication bias is based on the funnel plot, which compares intervention effects to weights, standard errors, or sample sizes [Macaskill 2001]. The typical thumbprint of publication bias is the presence of more favorable effects among small studies than among large studies. If increasing weights are plotted on the vertical axis and increasing ratios (RR or PR more favorable to the comparison condition) on the horizontal axis as in funnel plots 1.1, 1.2, 2.1, and 2.2, the bottom right portion of the figure (representing small studies with less favorable results) will contain fewer studies than the bottom left (representing small studies with more favorable results).

Funnel plots for interventions compared to minimal or no HIV prevention do not appear to indicate publication bias because the population of studies on the right side (indicating weaker effects) of the summary effect is at least dense as on the left side (indicating stronger effects) [Figure 2 and Figure 3].

Figure 2.   1.1 No tx cntrl RR.RR for interventions vs minimal to no HIV prevention

Review: Interventions to reduce risk for sexual transmission of HIV among men who have sex with men
Comparison: 01 Intervention vs minimal to no HIV prevention
Outcome: 01 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]
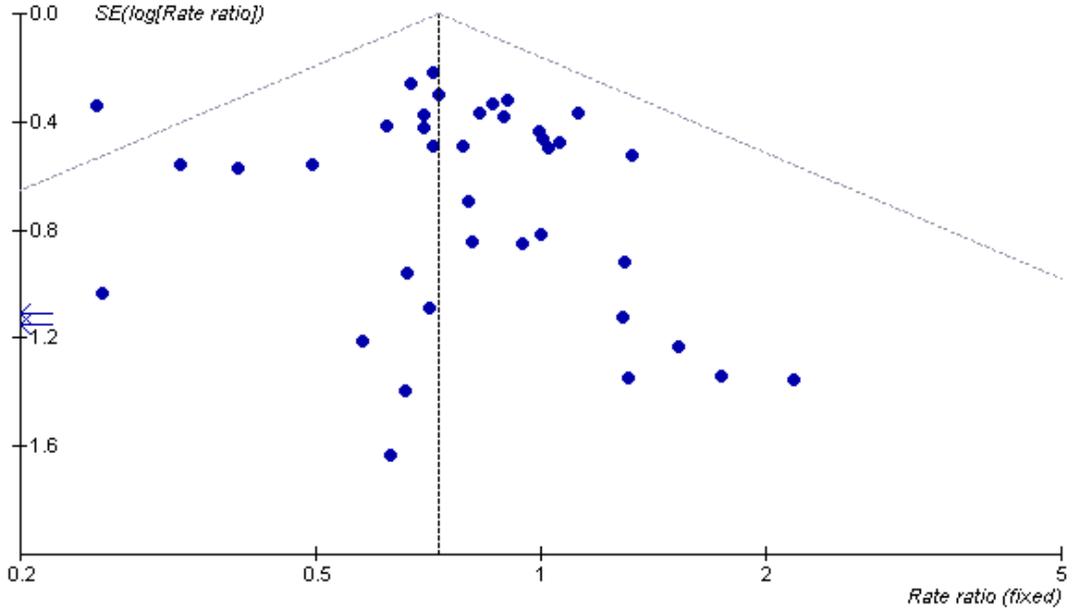
Figure 3. 1.2 No tx cntrl PR.PR for interventions vs minimal to no HIV prevention

Review: Interventions to reduce risk for sexual transmission of HIV among men who have sex with men
Comparison: 01 Intervention vs minimal to no HIV prevention
Outcome: 02 PR Proportion reporting any unprotected sex [# = RR substituted for PR]
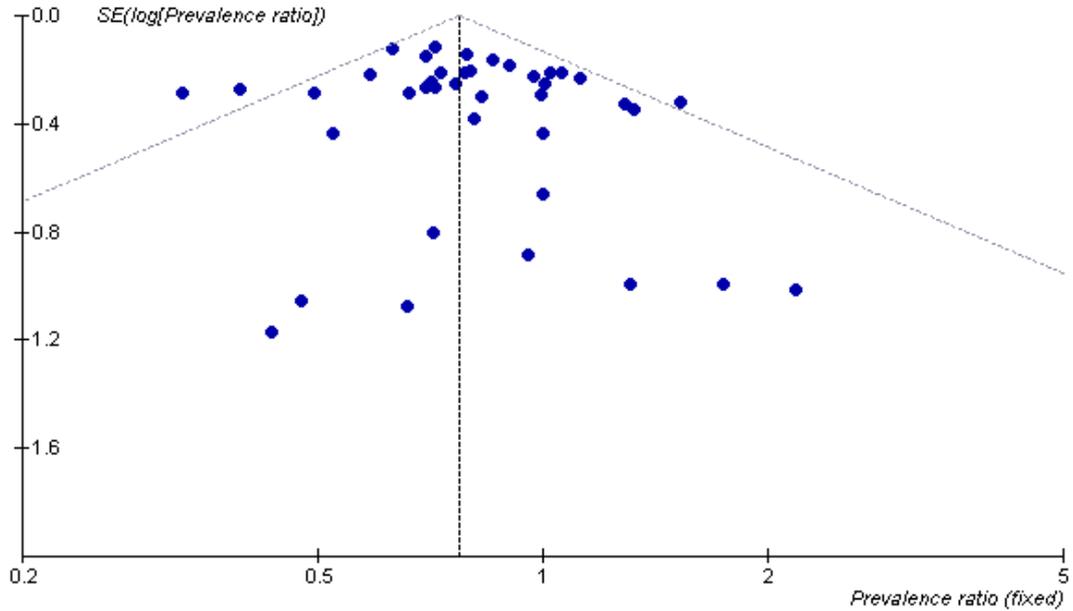
However in the case of interventions compared to standard or other HIV prevention, there may be somewhat fewer studies on the right side of the summary effect than on the left side [Figure 4 and Figure 5]. If this difference is due to publication bias, estimates from meta-analysis may tend to overstate effects for interventions compared to standard or other HIV prevention.

Figure 4. 2.1 Some tx cntrl RR.RR Experimental vs standard interventions

Review: Interventions to reduce risk for sexual transmission of HIV among men who have sex with men
Comparison: 02 Experimental vs standard or other HIV prevention
Outcome: 01 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]
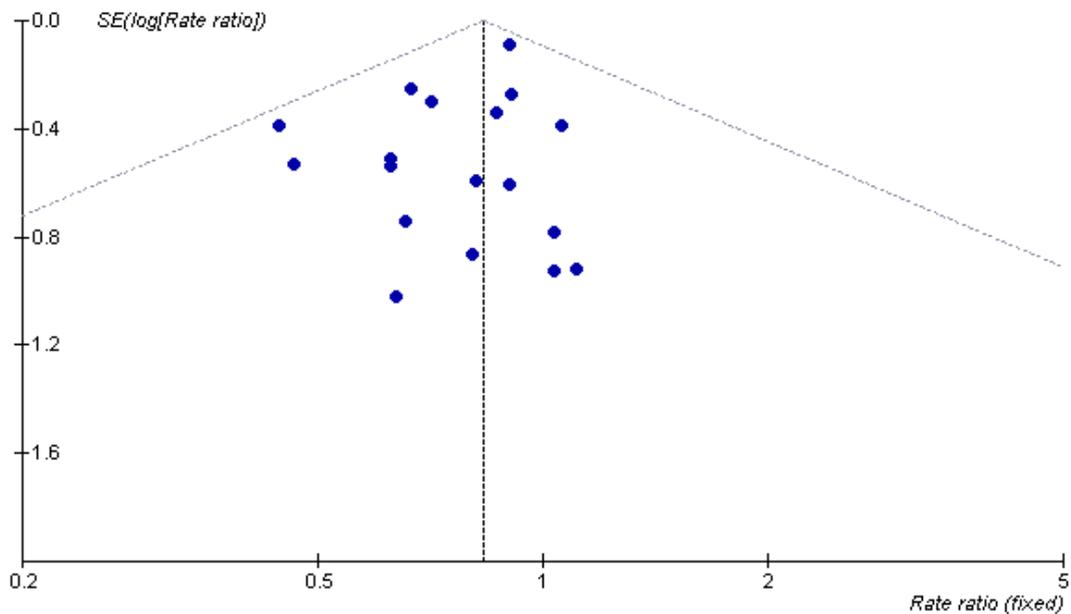
Figure 5.   2.2 Some tx cntrl PR.PR Experimental vs standard interventions

Review: Interventions to reduce risk for sexual transmission of HIV among men who have sex with men
Comparison: 02 Experimental vs standard or other HIV prevention
Outcome: 02 PR Proportion reporting any unprotected sex   [# = RR substituted for PR]
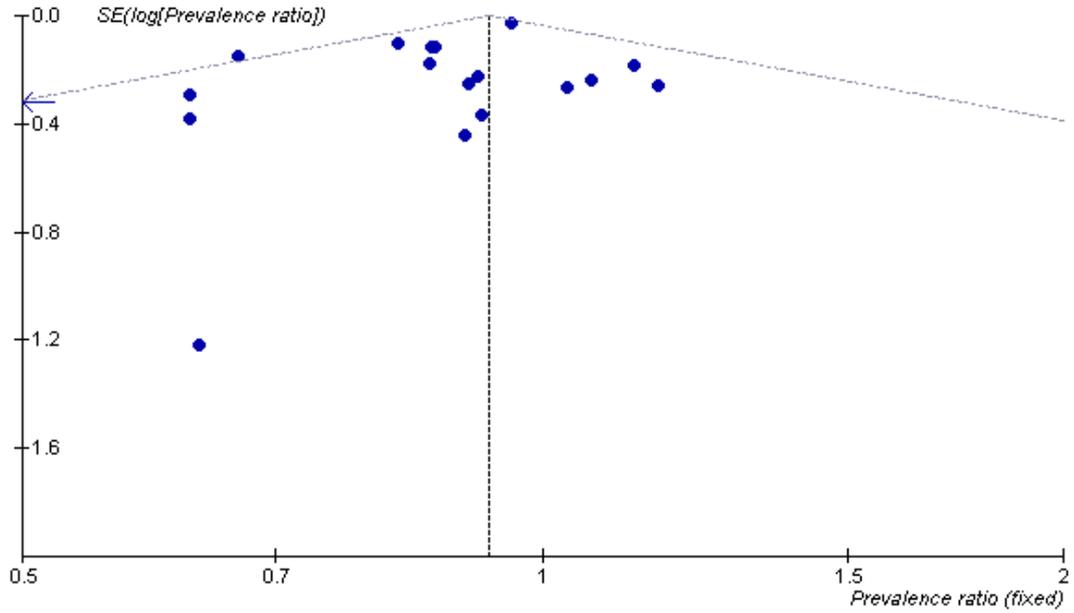
DISCUSSION

In randomized controlled trials and non-randomized trials with independent control conditions as described above, behavioral interventions for MSM reduced unprotected anal sex by 27% compared to minimal or no intervention, and reduced the proportion of men reporting any unprotected anal sex by 23%. When compared to standard or other HIV prevention, experimental interventions reduced number of episodes of or partners for unprotected anal sex by 17% and proportion reporting any unprotected anal sex by 7%. These effects were also evident in subgroup analyses of small group and community-level interventions, and somewhat less so for individual-level interventions.

The four characteristics most closely associated with effectiveness of small group and individual-level interventions in multivariable models warrant particular mention. As expected, effects were not as strong when interventions were contrasted against control conditions that included some content related to HIV prevention. Statistical significance results from the interplay of sample size, the strength of the intervention, and the strength of the comparison condition. When designing new studies, researchers should not underestimate the potential impact of active control conditions on behavior. Control for demand (e.g., changing behavior because that is what participants think is expected of them) and control for attention (e.g., changing behavior because participants are engaged in other activities or feel cared about during a research process) are important issues that should be considered when examining the marginal cost of delivering HIV prevention content. Studies in which the comparison condition receives some intervention relevant to HIV prevention will require larger sample sizes. But if they are successful, they will lead to greater confidence that the intervention is effective beyond attention and demand characteristics of the comparison condition.

Because studies with count outcomes showed the most favorable results, researchers and front-line prevention workers should be aware that some risk reduction comes in the form of fewer occasions of or partners for unprotected sex, even among those who do not completely eliminate their risk. A reduction in number of occasions of unprotected sex may have an important impact on HIV transmission rates, particularly if the number of partners for unprotected sex and the density of unprotected sexual networks also decrease.

The fact that the most favorable results were observed among studies with a shorter intervention span suggests that a clear and focused risk reduction message may be most effective, at least for MSM represented in these studies. Finally, effects were somewhat stronger (though not statistically significantly so) in studies with greater attrition in the comparison condition than in the intervention condition. Therefore, differential retention may be a threat to

validity in these studies, so extra effort should be given to retaining participants in comparison conditions.

Among community-level interventions, rate ratios of .60 or less, indicating a 40% or greater reduction in unprotected sex, were observed in studies with random assignment, shorter recall periods and longer follow-up times, those with more than 25% non-gay identifying MSM, those with more than 90% white participants, and interventions addressing development of personal skills.

We considered non-random assignment a potential threat to validity. If intervention effects had been stronger among studies with non-random assignment, their results could have been considered biased and a good argument could be raised for their exclusion from these meta-analyses. However intervention effects were actually stronger (although not statistically significantly so) in studies with random assignment. This result offers reassurance that inclusion of non-randomized trials (which still had to include baseline data for each condition and no evident source of bias in assignment) in this meta-analysis did not introduce a bias toward favorable effects.

It is encouraging to see effectiveness among community-level studies with longer follow-up times (6-16 months). Shorter recall periods may facilitate better recall and may improve chances of detecting an intervention effect. But in combination with a longer follow-up time, a short recall period implies an extended interim period for which no data are available. This problem can be addressed by collecting multiple waves of follow-up data.

The greater effectiveness of interventions that include efforts to promote personal skills such as keeping condoms readily available, avoiding excess intoxicants, self-reinforcement for behavior change, and behavioral self-management indicates that such content merits strong consideration in development and delivery of interventions for MSM. The greater effectiveness of interventions including more than 25% non-gay identifying MSM suggests that when they can be reached, these men are no less responsive than gay-identified men to risk reduction efforts.

The finding that interventions were most effective for majority white populations underscores the critical need for effective interventions for MSM of African and Latino descent. An adaptation of the Kelly 1991 Popular Opinion Leader Model has recently shown success with young African American MSM in a nonrandomized trial in North Carolina [Jones 2008].

Information on biologic outcomes was too limited to warrant quantitative analysis. In one study, a small group intervention for STD clinic patients showed modest reductions in unprotected sex but also resulted in more STD infections than a standard 1-on-1 counseling session about sexual risk behavior [Imrie 2001*@]. However an enhanced individual-level intervention with a similarly modest reduction in unprotected sex was accompanied by a substantial reduction in new HIV infections [Explore 2004*]. A

third study of a small group intervention for HIV-positive MSM found (nonsignificantly) fewer non-viral STDs in the enhanced intervention than in the standard intervention at followup, but baseline differences made the results difficult to interpret [Wolitski 2005*]. Many more trials are needed to measure the effects of behavioral interventions on HIV and STD incidence.

Some populations at high risk, particularly African-American [Leone 2004] and Latino MSM, and MSM in countries where English is not the primary language, have been underrepresented in intervention research. Factors affecting HIV risk are likely to differ among such populations [Courtenay-Quirk 2003; Coleman 2003; Zea 2003; Millett 2004]. Use of alcohol and drugs, particularly methamphetamine, and attending bathhouses, sex clubs, and circuit parties may be associated with risky sex among MSM [Patterson 2003 drug; Lister 2003; Crosby 2003; Semple 2003]. A wide range of effective interventions specifically designed for those at highest risk is urgently needed.

We chose unprotected anal sex as the outcome for these analyses because it is clearly identified as a risk factor for HIV transmission and it was available for all studies. But for better or worse, MSM are also using other behavioral strategies in an effort to reduce risk. Empirical examination of the effects of serosorting, negotiated safety, withdrawal before ejaculation, strategic positioning, and partner selection on HIV transmission is urgently needed [Hoff 2004]. Perception of partner's serostatus may be incorrect, so the effectiveness of such strategies in avoiding HIV infection is unknown. Availability of new treatments may contribute to complacency about HIV prevention [Demmer 2003]. In recent years the internet has become an important factor in the HIV epidemic [AIDS Alert 2004], but it may also be useful in prevention [AIDS Alert 2003] and in partner notification [CDC 2004]. These relatively new factors that may influence behaviors should be considered in future research.

AUTHORS' CONCLUSIONS

Implications for practice

Meta-analysis provides a comprehensive view of results of studies conducted to date. First, we conclude that behavioral HIV prevention interventions for MSM reduce self-reported unprotected sex and they should be funded.

The intervention content item associated with the greatest effectiveness among small group and individual-level interventions was a focus on losses rather than gains. It may be important in some settings not to shy away from discussion of losses associated with risky sexual behavior and HIV infection.

The intervention content item associated with the greatest effectiveness among community-level interventions was personal skills such as keeping condoms readily available, avoiding excess intoxi-

cants, self-reinforcement for behavior change, and behavioral self-management. Intervention curricula should address these skills, and staff should be trained in their delivery. Community-level interventions may also benefit from longer follow-up times.

The greater effectiveness of interventions including more than 25% non-gay identifying MSM suggests that when they can be reached, these men may be more responsive than gay-identified men to risk reduction efforts. Even though their initial level of risk behavior tends to be lower than that of gay-identified men, non-gay identified MSM may have less exposure to previous prevention messages, so their initial exposure may have a greater impact.

Implications for research

Results of previous studies provide a benchmark for expectations in new studies. Meta-analysis can inform future design and implementation in terms of target populations, sample size, settings, goals for process measures, and intervention content.

First and foremost, there is a critical need for effective interventions for MSM of African and Latino descent.

When effects differ by design variables, which are deliberately selected and planned, awareness of these characteristics may be beneficial to future designs. Given results from the multivariable model above, researchers designing new studies to measure intervention effects should strongly consider measuring unprotected sex as count outcomes including both the number of partners for unprotected sex and the number of occasions of unprotected sex. Researchers designing future small group and individual-level studies should also keep in mind that to date, effects of the greatest magnitude have been observed in studies that used a shorter intervention span (up to 1 month).

Among small group and individual-level studies, effects were also greatest when the comparison condition included minimal to no HIV prevention content. However statistically significant favorable effects were also seen when the comparison condition included standard or other HIV prevention content. Researchers choosing the latter option for new studies should plan for larger sample sizes based on the smaller expected net intervention effect noted above.

When effects differ by implementation variables, which become evident as the study is conducted but are not usually selected or planned, caution may be advised so that future studies can reduce bias. Because intervention effects were somewhat stronger (though not statistically significantly so) in studies with a greater attrition in the comparison condition, differential retention may be a threat to validity. Extra effort should be given to retaining participants in comparison conditions.

While community-level interventions may require longer follow-up times, measurement of their effects also benefits from shorter recall periods, which may necessitate multiple waves of data collection. Also among community-level interventions, intervention effects were strongest among studies with random assignment of groups or communities. Therefore the inclusion of studies where assignment of groups or communities was by convenience did not exaggerate the summary effect.

Further research measuring the incidence of HIV and other STDs is needed. Because most studies were conducted among mostly white men in the US and Europe, more evaluations of interventions are needed for African-American and Hispanic MSM and MSM in the developing world. More research is also needed to further clarify which behavioral strategies (e.g., reducing unprotected anal sex, having oral sex instead of anal sex, reducing number of partners, avoiding serodiscordant partners, strategic positioning, or reducing anal sex even with condom use) are most effective in reducing transmission among MSM, the messages most effective in promoting these behaviors, and the methods and settings in which these messages can be most effectively delivered.

REFERENCES

References to studies included in this review

Amirkhanian 2005@ *{published data only}*

* Amirkhanian YA, Kelly JA, Kabakchieva E, et al. A randomized social network HIV prevention trial with young men who have sex with men in Russia and Bulgaria. *AIDS* 2005;**19**:1897-1905.

Amirkhanian YA, Kelly JA, Kabakchieva E, McAuliffe TL, Vassileva S. Evaluation of a social network HIV prevention intervention program for young men who have sex with men in Russia and Bulgaria. *AIDS Education and Prevention* 2003;**15**(3):205–220.

Carballo-Diéguez2005 *{published and unpublished data}*

Carballo-Diéguez A, Dolezal C, Leu C-S, Nieves L, Díaz F, Decena C, Balan I. A randomized controlled trial to test an HIV-prevention intervention for Latino gay and bisexual men: Lessons learned. *AIDS Care* 2005;**17**:314–328.

CDC ACDP 1999@ *{published and unpublished data}*

CDC AIDS Community Demonstration Projects. Community-level HIV intervention in 5 cities: final outcome data from the CDC AIDS Community Demonstration Projects. *Am J Public Health* 1999;**89**:336–345.

Choi 1996@ *{published data only}*

Choi KH, Lew S, Vittinghoff E, et al. The efficacy of brief group counseling in HIV risk reduction among homosexual Asian and Pacific Islander men. *AIDS* 1996;**10**:81–87.

Cleary 1995*@ *{published data only}*

Cleary PD, Van Devanter N, Steilen M, Stuart A, Shipton-Levy R, McMullen W, et al. A randomized trial of an education and support program for HIV-infected individuals. *AIDS* 1995;**9**:1271–1278.

Coates 1989@ *{published and unpublished data}*

Coates TJ, McKusick L, Kuno R, et al. Stress reduction training changed number of sexual partners but not immune function in men with HIV. *Am J Public Health* 1989;**79**:885-7.

Dilley 2002 d*@ *{published data only (unpublished sought but not used)}*

Dilley JW, Woods WJ, Sabatino J, et al. Changing sexual behavior among gay male repeat testers. *J Acquir Immune Defic Syndr* 2002;**30**:177–186.

Dilley 2002 n*@ *{published data only (unpublished sought but not used)}*

Dilley JW, Woods WJ, Sabatino J, et al. Changing sexual behavior among gay male repeat testers. *J Acquir Immune Defic Syndr* 2002;**30**:177–186.

Dilley 2007*# *{published data only}*

Dilley JW, Woods WJ, Loeb L, Nelson K, Sheon N, Mullan J, Adler B, Chen S, McFarland W. Brief cognitive counseling with HIV testing to reduce sexual risk among men who have sex with men: Results from a randomized controlled trial using paraprofessional counselors. *J Acquir Immune Defic Syndr* 2007;**44**(5):569–577.

Elford 2001@ *{published and unpublished data}*

Elford J, Bolding G, Sherr L. Peer education has no significant impact on HIV risk behaviours among gay men in London. *AIDS* 2001;**15**(4):531–537.

Explore 2004* *{published and unpublished data}*

Chesney MA, Koblin BA, Barresi PJ, et al. An individually tailored intervention for HIV prevention: baseline data from the EXPLORE Study. *Am J Public Health* 2003;**93**(6):933–8.

EXPLORE Study Team. Effects of a behavioural intervention to reduce acquisition of HIV infection among men who have sex with men: the EXPLORE randomised controlled study. *Lancet* 2004;**364**:41–50.

Flowers 2002 *{published and unpublished data}*

Flowers P, Frankis J, Hart G. Evidence and the evaluation of a community-level intervention: research the Gay Men's Task Force. *Watson J, Platt S, eds. Researching Health Promotion.* London: Routledge, 2000:102–124.

* Flowers P, Hart GJ, Williamson LM, Frankis JS, Der GJ. Does bar-based, peer-led sexual health promotion have a community-level effect amongst gay men in Scotland?. *Intl J STD & AIDS* 2002;**13**: 102–108.

Hart GJ, Williamson LM. Increase in HIV sexual risk behaviour in homosexual men in Scotland, 1996-2002: prevention failure?. *Sex Transm Infect* 2005;**81**:367–372.

Williamson LM, Hart GJ, Flowers P, et al. The Gay Men's Task Force: the impact of peer education on the sexual health behaviour of homosexual men in Glasgow. *Sex Transm Infect* 1997;**75**:242–246.

Gold 1995 po* *{published data only (unpublished sought but not used)}*

Gold RS, Rosenthal DA. Preventing unprotected anal intercourse in gay men: A comparison of two intervention techniques. *Intl J STD AIDS* 1995;**6**:89–94.

Gold 1995 sj* *{published data only (unpublished sought but not used)}*

Gold RS, Rosenthal DA. Preventing unprotected anal intercourse in gay men: A comparison of two intervention techniques. *Intl J STD AIDS* 1995;**6**:89–94.

Gold 1998 rse *{published data only (unpublished sought but not used)}*

Gold RS, Rosenthal DA. Examining self-justifications for unsafe sex as a technique of AIDS education: the importance of personal relevance. *Intl J STD AIDS* 1998;**9**:208–213.

Gold 1998 sjp* *{published data only (unpublished sought but not used)}*

Gold RS, Rosenthal DA. Examining self-justifications for unsafe sex as a technique of AIDS education: the importance of personal relevance. *Intl J STD AIDS* 1998;**9**:208–213.

Harding 2004 *{published and unpublished data}*

Harding R, Bensley J, Corrigan N, Franks L, Stratman J, Waller Z, Warner J. Outcomes and lessons from a pilot RCT of a community-based HIV prevention multi-session group intervention for gay men. *AIDS Care* 2004;**16**:581–585.

Healthy Living 2007# *{published data only}*

Gore-Felton C, Rotheram-Borus MJ, Weinhardt LS, et al. The Healthy Living Project: an individually tailored, multidimensional intervention for HIV-infected persons. *AIDS Education & Prevention* 2005;**17**(1 Suppl A):21–39.

* Healthy Living Project Team. Effects of a behavioral intervention to reduce risk of transmission among people living with HIV: the healthy living project randomized controlled study. *J Acquir Immune Defic Syndr* 2007;**44**(2):213–21.

Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men (Review)   **25**

Copyright © 2009 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hoff 1997 *{unpublished data only}*

Hoff CC, Kegeles SM, Acree M, et al.Looking for men in all the wrong places. HIV prevention small-group programs do not reach high-risk gay men. *AIDS* 1997;**11**:829–31.

Imrie 2001*@ *{published data only}*

Imrie J, Stephenson JM, Cowan FM, et al.A cognitive behavioral intervention to reduce sexually transmitted infections among gay men: randomised trial. *BMJ* 2001;**322**:1451-6.

Kalichman 2001# *{published data only (unpublished sought but not used)}*

Kalichman SC, Rompa D, Cage M. Group intervention to reduce HIV transmission risk behavior among persons living with HIV/AIDS. *Behavior Modification* 2005;**29**(2):256–285.

* Kalichman SC, Rompa D, Cage M, et al.Effectiveness of an intervention to reduce HIV transmission risks in HIV-positive people. *Am J Prev Med* 2001;**21**(2):84–92.

Kegeles 1996@ *{published data only (unpublished sought but not used)}*

Kegeles SM, Hays RB, Coates TJ. The Mpowerment project: a community-level HIV prevention intervention for young gay men. *Am J Public Health* 1996;**86**:1129-1136.

Kegeles 2002@ *{published data only (unpublished sought but not used)}*

Kegeles SM, Rebchook GM, Hays RB, Pollack LM. Staving off increases in young gay/bisexual men's risk behavior in the HAART era. XIV International Conference on AIDS. Barcelona, 2002:Abstract # MoPeD3549.

Kelly 1989# *{published data only (unpublished sought but not used)}*

Kelly JA, St. Lawrence JS, Hood HV, et al.Behavioral intervention to reduce AIDS risk activities. *J Consulting Clin Psychol* 1989;**57**: 60–67.

Kelly 1991 *{published data only (unpublished sought but not used)}*

Kelly JA, St. Lawrence JS, Diaz YE, et al.HIV risk behavior reduction following intervention with key opinion leaders of population: an experimental analysis. *Am J Public Health* 1991;**81**:168–171.

Kelly 1993 cb *{published data only (unpublished sought but not used)}*

Kelly JA, Muphy DA, Bahr R, et al.Outcome of cognitive-behavioral and support group brief therapies for depressed, HIV-infected persons. *Am J Psychiatry* 1993;**150**(11):1679–1686.

Kelly 1993 ss *{published data only (unpublished sought but not used)}*

Kelly JA, Muphy DA, Bahr R, et al.Outcome of cognitive-behavioral and support group brief therapies for depressed, HIV-infected persons. *Am J Psychiatry* 1993;**150**(11):1679–1686.

Kelly 1997 *{published data only}*

Kelly JA, Murphy DA, Sikkema KJ, et al.Randomised, controlled, community-level HIV-prevention intervention for sexual-risk behaviour among homosexual men in US cities. *Lancet* 1997;**350**: 1500–05.

Miller 1998 *{published and unpublished data}*

Miller RL, Klotz D, Eckholdt HM. HIV prevention with male prostitutes and patrons of hustler bars: replication of an HIV preventive intervention. *American Journal of Community Psychology* 1998;**26** (1):97–131.

Patterson 2003 b# *{published data only (unpublished sought but not used)}*

Patterson TL, Shaw WS, Semple SJ. Reducing the sexual risk behaviors of HIV+ individuals: outcome of a randomized controlled trial. *Ann Behav Med* 2003;**25**:137–145.

Patterson 2003 c# *{published data only (unpublished sought but not used)}*

Patterson TL, Shaw WS, Semple SJ. Reducing the sexual risk behaviors of HIV+ individuals: outcome of a randomized controlled trial. *Ann Behav Med* 2003;**25**:137–145.

Patterson 2003 s# *{published data only (unpublished sought but not used)}*

Patterson TL, Shaw WS, Semple SJ. Reducing the sexual risk behaviors of HIV+ individuals: outcome of a randomized controlled trial. *Ann Behav Med* 2003;**25**:137–145.

Peterson 1996 1s *{published and unpublished data}*

Peterson JL, Coates TJ, Catania J, et al.Evaluation of an HIV risk reduction intervention among African-American homosexual and bisexual men. *AIDS* 1996;**10**:319-25.

Peterson 1996 3s *{published and unpublished data}*

Peterson JL, Coates TJ, Catania J, et al.Evaluation of an HIV risk reduction intervention among African-American homosexual and bisexual men. *AIDS* 1996;**10**:319-25.

Picciano 2001# *{published data only}*

Picciano JF, Roffman RA, Kalichman SC, et al.A telephone based brief intervention using motivational enhancement to facilitate HIV risk reduction among MSM: a pilot study. *AIDS and Behavior* 2001; **5**(3):251–262.

R-B 2004 in person@ *{published data only (unpublished sought but not used)}*

Rotheram-Borus MJ, Swendeman D, Comulada WS, Weiss RE, Lee M, Lightfoot M. Prevention for Substance-Using HIV-Positive Young People: Telephone and In-Person Delivery. *J Acquir Immune Defic Syndr* 2004;**37**:S68–S77.

R-B 2004 phone@ *{published data only (unpublished sought but not used)}*

Rotheram-Borus MJ, Swendeman D, Comulada WS, Weiss RE, Lee M, Lightfoot M. Prevention for Substance-Using HIV-Positive Young People: Telephone and In-Person Delivery. *J Acquir Immune Defic Syndr* 2004;**37**:S68–S77.

Read 2006*# *{published data only}*

Read SJ, Miller LC, Appleby PR, Nwosu ME, Reynaldo S, Lauren A, Putcha A. Socially Optimized Learning in a Virtual Environment: Reducing Risky Sexual Behavior Among Men Who Have Sex with Men. *Human Communication Research* 2006;**32**:1–34.

Richardson 2004 g@ *{published data only}*

Richardson JL, Milam J, McCutchan A, Stoyanoff S, Bolan R, Weiss J, et al.Effect of brief safer-sex counseling by medical providers to HIV-1 seropositive patients: a multi-clinic assessment. *AIDS* 2004; **18**:1179 -1186.

Richardson 2004 L@ *{published data only}*

Richardson JL, Milam J, McCutchan A, Stoyanoff S, Bolan R, Weiss J, et al.Effect of brief safer-sex counseling by medical providers to HIV-1 seropositive patients: a multi-clinic assessment. *AIDS* 2004; **18**:1179 -1186.

Roffman 1997 *{published and unpublished data}*

Roffman RA, Picciano JF, Ryan R, et al.HIV-prevention group counseling delivered by telephone: an efficacy trial with gay and bisexual men. *AIDS and Behavior* 1997;**1**(2):137–154.

Roffman 1998# *{published data only (unpublished sought but not used)}*

Roffman RA, Stephens RS, Curtin L, et al.Relapse prevention as an interventive model for HIV risk reduction in gay and bisexual men. *AIDS Ed Prevent* 1998;**10**(1):1–18.

Rosser 1990 c@ *{published data only}*
Rosser BRS. Evaluation of the efficacy of AIDS education interventions for homosexually active men. *Health Education Research* 1990; **5**(3):299–308.

Rosser 1990 e@ *{published data only}*
Rosser BRS. Evaluation of the efficacy of AIDS education interventions for homosexually active men. *Health Education Research* 1990; **5**(3):299–308.

Rosser 1990 s@ *{published data only}*
Rosser BRS. Evaluation of the efficacy of AIDS education interventions for homosexually active men. *Health Education Research* 1990; **5**(3):299–308.

Rosser 1990 v@ *{published data only}*
Rosser BRS. Evaluation of the efficacy of AIDS education interventions for homosexually active men. *Health Education Research* 1990; **5**(3):299–308.

Rosser 2002*@ *{published data only}*
Rosser BRS, Bockting WO, Rugg DL, et al.A randomized controlled intervention trial of a sexual health approach to long-term HIV risk reduction in men who have sex with men. Effects of the intervention on unsafe sexual behavior. *AIDS Ed Prevent* 2002;**14**(Suppl 3):59–71.

Rotheram-Borus 2001@ *{published data only (unpublished sought but not used)}*
Rotheram-Borus MJ, Lee MB, Murphy DA, Futterman D, Duan N, Birnbaum JM, Lightfoot M, the Teens Linked to Care Consortium. Efficacy of a Preventive Intervention for Youths Living With HIV. *Am J Public Health* 2001;**91**:400–405.

Sampaio 2002*@ *{published and unpublished data}*
Sampaio M, Brites C, Stall R, Hudes ES, Hearst N. Reducing AIDS risk among men who have sex with men in Salvador, Brazil. *AIDS and Behavior* 2002;**6**(2):173–181.

Shepherd 1997 *{published and unpublished data}*
Shepherd J, Weare K, Turner G. Peer-led sexual health promotion with young gay and bisexual men: results of The HAPEER Project. *Health Education* 1997;**97**:204–212.

Shoptaw 2005 cbt* *{published and unpublished data}*
Shoptaw S, Reback CJ, Peck JA, Yang X, Rotheram-Fuller E, Larkins S, et al.Behavioral treatment approaches for methamphetamine dependence and HIV-related sexual risk behaviors among urban gay and bisexual men. *Drug and Alcohol Dependence* 2005 May 9;**78**(2):125–34.

Shoptaw 2005 g* *{published and unpublished data}*
Shoptaw S, Reback CJ, Peck JA, Yang X, Rotheram-Fuller E, Larkins S, et al.Behavioral treatment approaches for methamphetamine dependence and HIV-related sexual risk behaviors among urban gay and bisexual men. *Drug and Alcohol Dependence* 2005 May 9;**78**(2):125–34.

Sorensen 2003* *{published and unpublished data}*
Sorensen JL, Dilley J, London J, Okin RL, Delucchi KL, Phibbs CS. Case management for substance abusers with HIV/AIDS: a randomized clinical trial. *American Journal of Drug and Alcohol Abuse* 2003; **29**:133–150.

Stall 1999@ *{published data only}*
Stall RD, Paul JP, Barrett DC, Crosby GM, Bein E. An outcome evaluation to measure changes in sexual risk-taking among gay men undergoing substance use disorder treatment. *Journal of Studies on Alcohol* 1999;**60**(6):837–845.

Tudiver 1992 c4@ *{published data only}*
Tudiver F, Myers T, Kurtz RG, et al.The Talking Sex Project. *Evaluation and the Health Professions* 1992;**15**(4):26–42.

Tudiver 1992 p1@ *{published data only}*
Tudiver F, Myers T, Kurtz RG, et al.The Talking Sex Project. *Evaluation and the Health Professions* 1992;**15**(4):26–42.

Valdiserri 1989* *{published and unpublished data}*
Valdiserri RO, Lyter DW, Leviton LC, et al.AIDS prevention in homosexual and bisexual men: results of a randomized trial evaluating two risk reduction interventions. *AIDS* 1989;**3**:21–26.

Wolitski 2005* *{published and unpublished data}*
Hoff CC, Pals SL, Purcell DW, Parsons JT, Halkitis PN, Remien RH, Gomez C. Examining the role of partner status in an HIV prevention trial targeting HIV-positive gay and bisexual men: the seropositive urban men's intervention trial (SUMIT). *AIDS and Behavior* 2006; **10**:637–648.
* Wolitski RJ, Gómez CA, Parsons JT, and the SUMIT Study Group. Effects of a peer-led behavioral intervention to reduce HIV transmission and promote serostatus disclosure among HIV-seropositive gay and bisexual men. *AIDS* 2005;**19(Suppl 1)**:S99-S109.

References to studies excluded from this review

Amirkhanian 2003 *{published data only}*
Amirkhanian YA, Kelly JA, Kabakchieva E, McAuliffe TL, Vassileva S. Evaluation of a social network HIV prevention intervention program for young men who have sex with men in Russia and Bulgaria. *AIDS Education and Prevention* 2003;**15**(3):205–220.

Antoni 2000 *{published data only}*
Antoni MH, Cruess S, Cruess DG, et al.Cognitive-behavioral stress management reduces distress and 24-hour urinary free cortisol output among symptomatic HIV-infected gay men. *Annals of Behavioral Medicine* 2000;**22**(1):29–37.

Blake 2001 *{published data only}*
Blake SM, Ledsky R, Lehman T, Goodenow C, Sawyer R, Hack T. Preventing sexual risk behaviors among gay, lesbian, and bisexual adolescents: the benefits of gay-sensitive HIV instruction in schools. *Am J Public Health* 2001;**91**(6):940–6.

Bowen 2007 *{published data only}*
Bowen AM, Horvath K, Williams ML. A randomized control trial of Internet-delivered HIV prevention targeting rural MSM. *Health Education Research* 2007;**22**(1):120–127.

Card 2001 *{published data only}*
Card JJ, Benner T, Shields JP, Feinstein N. The HIV/AIDS Prevention Program Archive (HAPPA): a collection of promising prevention programs in a box. *AIDS Education & Prevention* 2001;**13**(1):1–28.

Chesney 2003 *{published data only}*
Chesney MA, Chambers DB, Taylor JM, Johnson LM, Folkman S. Coping effectiveness training for men living with HIV: results from a randomized clinical trial testing a group-based intervention. *Psychosomatic Medicine* 2003;**65**:1038–1046.

Cote & Pepler 2002 *{published data only}*
Cote J, Pepler C. A randomized trial of a cognitive coping intervention for acutely ill HIV-positive men. *Nursing Research* 2002;**51**: 237-244.

Cote & Pepler 2005 *{published data only}*
Cote JK, Pepler C. Cognitive coping intervention for acutely ill HIV-positive men. *Journal of Clinical Nursing* 2005;**14**:321–326.

Cruess 2002 *{published data only}*
* Cruess S, Antoni MH, Hayes A, et al.Changes in Mood and Depressive Symptoms and Related Change Processes During Cognitive-Behavioral Stress Management in HIV-Infected Men. *Cognitive Therapy and Research* 2002;**26**(3):373–392.

DiIorio 2007 *{published data only}*
DiIorio C, McCarty F, Resnicow K, Lehr S, Denzmore P. REAL Men: a group-randomized trial of an HIV prevention intervention for adolescent boys. *Am J Public Health* 2007;**97**:1084–1089. [: doi: 10.2105/AJPH.2005.073411]

Fisher 2006 *{published data only}*
* Fisher JD, Fisher WA, Cornman DH, Amico RK, Bryan A, Friedland GH. Clinician-delivered intervention during routine clinical care reduces unprotected sexual behavior among HIV-infected patients. *J Acquir Immune Defic Syndr* 2006;**41**(1):44–52.

French 2000 *{published data only}*
French R, Power R, Mitchell S. An evaluation of peer-led STD/HIV prevention work in a public sex environment. *AIDS Care* 2000 Apr; **12**(2):225–234.

Heckman 2004 *{published data only}*
Heckman TG, Anderson ES, Sikkema KJ, Kochman A, Kalichman SC. Emotional distress in nonmetropolitan persons living with HIV disease enrolled in a telephone-delivered, coping improvement group intervention. *Health Psychology* 2004;**23**(1):94–100.

Hospers 1999 *{published data only}*
Hospers HJ, Debets W, Ross MW, Kok G. Evaluation of an HIV prevention intervention for men who have sex with men at cruising areas in the Netherlands. *AIDS and Behavior* 1999;**3**(4):359–366.

Huebner 2002 *{published data only}*
Huebner DM, Davis MC, Nemeroff CJ, Aiken LS. The impact of internalized homophobia on HIV preventive interventions. *Am J Comm Psych* 2002;**30**(3):327–48.

Jones 2008 *{published data only}*
* Jones KT, Gray P, Whiteside TO, Wang T, Bost D, Dunbar E, Foust E, Johnson WD. Evaluation of an HIV Prevention Intervention Adapted for Black Men Who Have Sex With Men. *Am J Public Health* 2008;**98**:1043–1050.

Kelly 1990 *{published data only}*
Kelly JA, St Lawrence JS, Betts R, Brasfield TL, Hood HV. A skills-training group intervention model to assist persons in reducing risk behaviors for HIV infection. *AIDS Education & Prevention* 1990;**2** (1):24–35.

Miller 1995 *{published and unpublished data}*
Miller RL. Assisting gay men to maintain safer sex: an evaluation of an AIDS service organization's safer sex maintenance program. *AIDS Educ Prevent* 1995;**7**(**Suppl**):48-63.

Nokes 2003 *{published data only}*
Nokes KM, Chew L, Altman C. Using a telephone support group for HIV-positive persons aged 50+ to increase social support and health-related knowledge. *AIDS Patient Care and STDs* 2003;**17**(7): 345–351.

Perry 1991 *{published data only}*
Perry S, Fishman B, Jacobsberg L, et al.Effectiveness of psychoeducational interventions in reducing emotional distress after human immunodeficiency virus antibody testing. *Arch Gen Psychiatry* 1991; **48**:143–7.

Reback 2004 *{published data only}*
* Reback CJ, Larkins S, Shoptaw S. Changes in the meaning of sexual risk behaviors among gay and bisexual male methamphetamine abusers before and after drug treatment. *AIDS and Behavior* 2004; **8**:87–98.

Remafedi 1994 *{published data only}*
Remafedi G. Cognitive and behavioral adaptations to HIV/AIDS among gay and bisexual adolescents. *Journal of Adolescent Health* 1994;**15**(2):142–148.

Rosser 1991 *{published data only}*
Rosser BRS. The Effects of Using Fear in Public AIDS Education on the Behaviour of Homosexually Active Men. *Journal of Psychology & Human Sexuality* 1991;**4**(3):123–134.

Toro-Alfonso *{published data only}*
Toro-Alfonso J, Varas-Diaz N, Andujar-Bello I. Evaluation of an HIV/AIDS prevention intervention targeting Latino gay men and men who have sex with men in Puerto Rico. *AIDS Education and Prevention* 2002;**14**(6):445–56.

Ziersch 2000 *{published data only}*
Ziersch A, Gaffney J, Tomlinson DR. STI prevention and the male sex industry in London: evaluating a pilot peer education programme. *Sex Transm Inf* 2000;**76**:447–453.

Zimmerman 1997 *{published data only}*
Zimmerman MA, Ramirez-Valles J, Suarez E, de la Rosa G, Castro MA. An HIV/AIDS prevention project for Mexican homosexual men: an empowerment approach. *Health Education & Behavior* 1997;**24**(2):177–90.

References to studies awaiting assessment

Antunes 1997 *{published data only}*
Antunes MC, Stall RD, Paiva V, et al.Evaluating an AIDS sexual risk reduction program for young adults in public night schools in Sao Paulo, Brazil. *AIDS* 1997;**11**(Suppl 1):S121–7.

CITY *{published data only}*
Guenther-Grey CA, Varnell S, Weiser JI, et al.Trends in Sexual Risk-Taking among Urban Young Men Who Have Sex with Men, 1999-2002. *Journal of the National Medical Association* 2005;**97**(7):38S–43S.

DiFranceisco 1998 *{unpublished data only}*
DiFranceisco W, Kelly JA, Sikkema KJ, et al.Differences between completers and early dropouts from 2 HIV intervention trials: A health belief approach to understanding prevention program attrition. *Am J Public Health* 1998;**88**(7):1068–1073.

Goodkin 1999 *{published data only}*
Goodkin K, Blaney NT, Feaster DJ, Baldewicz T, Burkhalter JE, Leeds B. A Randomized Controlled Clinical Trial of a Bereavement

Support Group Intervention in Human Immunodeficiency Virus Type 1-Seropositive and -Seronegative Homosexual Men. *Archives of General Psychiatry* 1999;**56**(1):52–59.

Mausbach 2007# *{published data only}*
* Mausbach BT, Semple SJ, Strathdee SA, Zians J, Patterson TL. Efficacy of a behavioral intervention for increasing safer sex behaviors in HIV-positive MSM methamphetamine users: Results from the EDGE study. *Drug and Alcohol Dependence* 2007;**87**(2-3):249–57.

NIMH 1998 *{published data only}*
The National Institute of Mental Health (NIMH) Multisite HIV Prevention Trial Group. The NIMH Multisite HIV Prevention Trial: Reducing HIV Sexual Risk Behavior. *Science* 1998;**280**(5371): 1889–1894.

Orellana 2006 *{published data only}*
Orellana ER, Picciano JF, Roffman RA, Swanson F, Kalichman SC. Correlates of nonparticipation in an HIV prevention program for MSM. *AIDS Education and Prevention* 2006;**18**(4):384–361.

Picciano 2007 *{published data only}*
Picciano JF, Roffman RA, Kalichman SC, Walker DD. Lowering obstacles to HIV prevention services: Effects of a brief telephone-based intervention using motivational enhancement therapy. *Ann Behav Med* 2007;**34**(2):177–187.

Remien *{published data only}*
Carballo-Dieguez A, Remien RH, Dolezal C, Wagner G. Unsafe sex in the primary relationships of Puerto Rican men who have sex with men. *AIDS and Behavior* 1997;**1**:9–17.
Remien RH. Couples of mixed HIV status: Challenges and strategies for intervention with couples. *L. Wicks (ed.), Psychotherapy and AIDS*. Washington, DC: Taylor & Francis, 1997.
Wagner G, Remien RH, Carballo-Dieguez A. "Extramarital" sex: is there an increased risk for HIV transmission? a study of male couples of mixed HIV status. *AIDS Education and Prevention* 1998;**10**(3): 245–256.

Smart Sex Quest 2004 *{published data only}*
Salyers Bull S, Llloyd L, Rietmeijer C, McFarlane M. Recruitment and retention of an online sample for an HIV prevention intervention targeting men who have sex with men: the Smart Sex Quest Project. *AIDS Care* 2004;**16**(8):931–943.

van Kesteren 2006 *{published data only}*
van Kesteren NMC, Kok G, Hospers HJ, Schippers J, de Wildt W. Systematic development of a self-help and motivational enhancement intervention to promote sexual health in HIV-positive men who have sex with men. *AIDS Patient Care and STDs* 2006;**20**(12):858–875.

Williams 2006 *{published data only}*
Williams ML, Bowen AM, Timpson SC, Ross MW, Atkinson JS. HIV prevention and street-based male sex workers: an evaluation of brief interventions. *AIDS Education and Prevention* 2006;**18**(3): 204–215.

Additional references

AIDS Alert 2003
Anonymous. Interventions can ease dangers of cybersex. Counselors educate while in chatrooms. *AIDS Alert* 2003;**18**:124–127.

AIDS Alert 2004
Anonymous. Study confirms the link between MSM and web. San Francisco officials use on-line interventions. *AIDS Alert* 2004;**19**: 19–20.

Baral 2007
Baral S, Sifakis F, Cleghorn F, Beyrer C. Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000-2006: a systematic review. *PLoS Medicine* 2007;**4** (12):e339. [: doi:10.1371/journal.pmed.0040339]

Bartholow 2006
Bartholow BN, Goli V, Ackers M, McLellan E, Gurwith M, Durham M, Greenberg AE. Demographic and behavioral contextual risk groups among men who have sex with men participating in a phase 3 HIV vaccine efficacy trial: implications for HIV prevention and behavioral/biomedical intervention trials. *J Acquir Immune Defic Syndr* 2006;**43**:594–602.

Becker 1998
Becker BJ. Synthesizing standardized mean-change measures. *Br J Math Stat Psychol* 1998;**41**:257–278.

Binson 1995
Binson D, Michaels S, Stall R, Coates TJ, Gagnon JH, Catania JA. Prevalence and social distribution of men who have sex with men: United States and its urban centers. *J Sex Research* 1995;**32**(3):245–254.

Bulterys 1997
Bulterys M, Morgenstern H, Weed DL. Quantifying the expected versus potential impact of a risk-factor intervention program. *Am J Public Health* 1997;**87**:867-8.

Calzavara 2000
Calzavara L, Burchell A, Major C, et al.Increasing HIV incidence among MSM repeat testers in Ontario, Canada, 1992-1998. XIII International AIDS Conference. Durban, South Africa, 2000. [: abstract ThOrC718]

Catania 2000
Catania JA, Morin SF, Canchola J, et al.U.S. priorities-HIV prevention. *Science* 2000;**290**:717.

CDC 1999a
Centers for Disease Control and Prevention. HIV/AIDS Surveillance Report, 1999. *HIV/AIDS Surveillance Report, 1999* 1999;**11**(2):32–33.

CDC 1999b
Centers for Disease Control and Prevention. [Resurgent bacterial sexually transmitted disease in men who have sex with men - King County, Washington, 1997-1999]. *MMWR* 1999;**48**:773-7.

CDC 2000
Centers for Disease Control and Prevention. HIV/AIDS Surveillance Report. 2000;**12**(2):table 6. [: http://www.cdc.gov/hiv/stats/hasr1202/table6.htm]

CDC 2001
Centers for Disease Control and Prevention. HIV/AIDS Surveillance Report. 2001;**13**(1).

CDC 2003
Centers for Disease Control and Prevention. HIV/AIDS Surveillance Report. 2003;**15**.

CDC 2004

Centers for Disease Control and Prevention (CDC). Using the Internet for partner notification of sexually transmitted diseases--Los Angeles County, California, 2003. *MMWR Morbidity & Mortality Weekly Report* 2004;**53**:129–31.

CDC 2007

Centers for Disease Control and Prevention. HIV/AIDS Surveillance Report for 2005, revised June 2007. 2007; **17, revised edition**:tables 1, 3, 19, 20. [: http://www.cdc.gov/hiv/topics/surveillance/resources/reports/.]

Chinn 2000

Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine* 2000;**19**:3127–31.

Choi 1994

Choi KH, Coates TJ. Prevention of HIV infection. *AIDS* 1994;**8**: 1371-89.

Cochrane Reviewers' Handbook

Deeks J, et al.Cochrane Reviewers' Handbook 4.2.2.

Cohen 2007

Cohen J. AIDS research: Promising AIDS vaccine's failure leaves field reeling Cohen J. AIDS research. Promising AIDS vaccine's failure leaves field reeling. [News] Science. 318(5847):28-9, 2007 Oct 5.. *Science* 2007;**318**(5847):28–9.

Coleman 2003

Coleman CL. Determinants of HIV and AIDS among young African-American men who have sex with men: a public health perspective. *Journal of National Black Nurses Association* 2003;**14**:25–29.

Conant 2004

Conant M. Losing the war on AIDS. *Journal of the American Academy of Dermatology* 2004;**51**(1 Suppl):S47–48.

Cook & Campbell

Cook TD, Campbell DT. Chapter 3: Quasi-Experiments: Nonequivalent Control Group Designs. *Quasi-Experimentation: Design & Analysis Issues for Field Settings*. Boston: Houghton Mifflin Company, 1979:95–146.

Courtenay-Quirk 2003

Courtenay-Quirk C, Wolitski RJ, Hoff C, Parsons JT. Interests in HIV prevention topics of HIV-seropositive men who have sex with men. *AIDS Educ Prevent* 2003;**15**:401–412.

Cox 2004

Cox J, Beauchemin J, Allard R. HIV status of sexual partners is more important than antiretroviral treatment related perceptions for risk taking by HIV positive MSM in Montreal, Canada. *Sexually Transmitted Infections* 2004;**80**:518–523.

Crepaz 2006

Crepaz N, Lyles CM, Wolitski RJ, et al.Do prevention interventions reduce HIV risk behaviours among people living with HIV? A meta-analytic review of controlled trials. *AIDS* 2006;**20**:143-157.

Crosby 2003

Crosby R, DiClemente RJ, Mettey A. Correlates of recent unprotected anal sex among men having sex with men attending a large sex resort in the South. *Sexually Transmitted Diseases* 2003;**30**:909–913.

Crosby 2006

Crosby RA, Holtgrave DR. Will sexual risk behaviour increase after being vaccinated for AIDS?. *International Journal of STD & AIDS* 2006;**17**:180–184.

Dawson 1996

Dawson C, Hartfield K. Developing a cost-effective media campaign addressing unprotected anal sex among gay men. *AIDS Educ Prevent* 1996;**8**:285-93.

Deeks 2002

Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine* 2002;**21**:1575–1600.

Demmer 2003

Demmer C. HIV prevention in the era of new treatments. *Health Promotion Practice* 2003;**4**:449–456.

DerSimonian 1986

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**:177–88.

Diaz 1996

Diaz RM, Stall RD, Hoff C, Diagle D, Coates TJ. HIV risk among Latino gay men in the Southwestern United States. *AIDS Education and Prevention* 1996;**8**:415–429.

Dillon 2000

Dillon B, Hecht FM, Swanson M, et al.Primary HIV infections associated with oral transmission. 7th Conference on Retrovirology and Opportunistic Infection. San Francisco, 2000. [: abstract no. 473]

Dukers 2007

Dukers NHTM, Fennema HSA, van der Snoek EM, et al.HIV incidence and HIV testing behavior in men who have sex with men: using three incidence sources, The Netherlands, 1984-2005. *AIDS* 2007;**21**:491–499.

Edwards 1966

Edwards JH. Some taxonomic implications of a curious feature of the bivariate normal surface. *Br J Prevention Social Med* 1966;**20**:42.

Egger 1997

Egger M, Davey Smith G, Schneider M, et al.Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**:629-34.

Ekstrand 1999

Ekstrand ML, Stall RD, Paul JP, et al.Gay men report high rates of unprotected anal sex with partners of unknown or discordant HIV status. *AIDS* 1999;**13**:1525-33.

Engels 2000

Engels EA, Schmid CH, Terrin I, et al.Heterogeneity and statistic significnace in meta-analysis: an emprical study of 125 meta-analyses. *Statist Med* 2000;**19**:1707–28.

Fisher HH 2006

Fisher HH, Purcell DW, Hoff CC, Parsons JT, O'Leary A. Recruitment source and behavioral risk patterns of HIV-positive men who have sex with men. *AIDS and Behavior* 2006;**10**:553–561.

Fisher JD 1992

Fisher JD, Fisher WA. Changing AIDS-risk behavior. *Psychol Bull* 1992;**111**:455-74.

Fisher JD 2006a

Fisher JD, Fisher WA, Amico KR, Harman JJ. An information-motivation-behavioral skills model of adherence to antiretroviral therapy. *Health Psychology* 2006;**25**(4):462–73.

Fisher JD 2006b

Fisher JD, Fisher WA, Cornman DH, et al.Clinician-delivered intervention during routine clinical care reduces unprotected sexual behavior among HIV-infected patients. *J Acquir Immune Defic Syndr* 2006;**41**:44–52.

Garber 2004

Garber DA, Silvestri G, Feinberg MB. Prospects for an AIDS vaccine: three big questions, no easy answers. *Lancet Infectious Diseases* 2004; **4**:397–413.

Gray 2003

Gray RH, Li X, Wawer MJ, Gange SJ, Serwadda D, Sewankambo NK, et al.Stochastic simulation of the impact of antiretroviral therapy and HIV vaccines on HIV transmission; Rakai, Uganda. *AIDS* 2003; **17**:1941–51.

Guzman 2003

Guzman R, Colfax GN, Wheeler S, Mansergh G, Marks G, Rader M, Buchbinder S. Negotiated safety relationships among a diverse sample of HIV-negative men who have sex with men. *J Acquir Immune Defic Syndr* 2005;**38**:82–86.

Hannan 1994

Hannan PJ, Murray D, Jacobs D, et al.Parameters to aid in the design and analysis of community trials: Intra-class correlations from the Minnesota Heart Health Program. *Epidemiology* 1994;**5**:88-95.

Harawa 2004

Harawa NT, Greenland S, Bingham TA, et al.Associations of race/ethnicity with HIV prevalence and HIV-related behaviors among young men who have sex with men in 7 urban centers in the United States. *J Acquir Immune Defic Syndr* 2004;**35**:526–536.

Harrison 1999

Harrison LH, do Lago RF, Friedman RK, et al.Incident HIV infection in a high-risk, homosexual, male cohort in Rio de Janeiro, Brazil. *J Acquir Immune Defic Syndr* 1999;**21**:408-12.

Hasselblad 1995

Hasselblad V, Hedges LV. Meta-analysis of screening and diagnostic tests. *Psych Bulletin* 1995;**117**:167-78.

Hawkins 2001

Hawkins DA. Oral sex and HIV transmission. *Sex Transm Infect* 2001;**77**:307-8.

Hedges 1994

Hedges LV. Fixed effects models. In: Cooper H, Hedges LV editor (s). *The handbook of research synthesis*. New York: Russell Sage Foundation, 1994:285-99.

Hedges 1998

Hedges LV, Vevea Jl. Fixed- and random-effects models in meta-analysis. *Psychol Meth* 1998;**3**:486-504.

Herbst 2005

Herbst JH, Sherba RT, Crepaz N, DeLuca JB, Zohrabyan L, Stall RD, et al.A meta-analytic review of HIV behavioral interventions for reducing sexual risk behavior of men who have sex with men for reducing sexual risk behavior. *J Acquir Immune Defic Syndr* 2005;**39**: 228-241.

Herbst 2007

Herbst JH, Beeker C, Mathew A, et al.The effectiveness of individual-, group-, and community-level HIV behavioral risk-reduction interventions for adult men who have sex with men. *Am J Prev Med* 2007;**32**(4S):S38–S67.

Hoff 2004

Hoff CC, Faigeles B, Wolitski RJ, Purcell DW, Gomez C, Parsons JT. Sexual risk of HIV transmission is missed by traditional methods of data collection. *AIDS* 2004;**18**:340–342.

Holmberg 1996

Holmberg SD. The estimated prevalence and incidence of HIV in 96 large U.S. metropolitan areas. *Am J Public Health* 1996;**86**:642-54.

Holtgrave 1995

Holtgrave DR, Qualls NL, Curran JW, et al.An overview of the effectiveness and efficiency of HIV prevention programs. *Public Health Rep* 1995;**110**:134–46.

Holtgrave 2000

Holtgrave DR, Pinkerton SD. Economics of HIV prevention. In: Peterson JL, DiClemente RJ editor(s). *Handbook of HIV prevention*. New York: Kluwer Academic/Plenum Publishers, 2000:286-96.

Holtgrave 2006

Holtgrave DR, Curran JW. What works, and what remains to be done, in HIV prevention in the United States. *Annu Rev Public Health* 2006;**27**:1–15.

Holtgrave 2007

Holtgrave DR, Kates J. HIV incidence and CDC's HIV prevention budget: an exploratory correlational analysis. *Am J Prev Med* 2007 Jan;**32**(1):63–7.

Imrie 2001

Imrie J, Stephenson JM, Cowan FM, et al.A cognitive behavioral intervention to reduce sexually transmitted infections among gay men: randomised trial. *BMJ* 2001;**322**:1451-6.

Johnson 2000

Johnson BT, Eagly AH. Quantitative synthesis of social psychologic research. In: Reis HT, Judd CM editor(s). *Handbook of research methods in social and personality psychology*. London: Cambridge University Press, 2000:496–528.

Johnson 2002b

Johnson WD, Semaan S, Hedges LV, et al.A protocol for the analytical aspects of a systematic review of HIV prevention research. *J Acquir Immune Defic Syndr* 2002;**30**(Suppl 1):S62–72.

Johnson 2004

Johnson WD, Flanders WD, Hill AN, McClellan WM, Goodman M, Holtgrave DR. Some new solutions for meta-analysis. PhD Work In Progress Day. Atlanta GA: Emory University School of Public Health, 2004 (Aug 26).

Johnson BT 2006

Johnson BT, Carey MP, Chaudoir SR, Reid AE. Sexual risk reduction for persons living with HIV: research synthesis of randomized controlled trials, 1993 to 2004. *J Acquir Immune Defic Syndr* 2006; **41**(5):642–650.

Johnston 2001

Johnston MI, Flores J. Progress in HIV vaccine development. *Curr Opin Pharmacol* 2001;**1**:504–10.

Kahn 2001

Kahn JG, Kegeles SM, Hays R, et al.Cost-effectiveness of the Mpowerment Project, a community-level intervention for young gay men. *J Acquir Immune Defic Syndr* 2001;**27**:482-91.

Kalichman 1996

Kalichman SC, Carey MP, Johnson BT. Prevention of sexually transmitted HIV infection: a meta-analytic review of the behavioral outcome literature. *Ann Behav Med* 1996;**18**:6–15.

Karon 1996

Karon JM, et al.Prevalence of HIV infection in the United States, 1984 to 1992. *JAMA* 1996;**276**(2):126–31.

Kegeles 1998

Kegeles SM, Hart GJ. Recent HIV-prevention interventions for gay men: individual, small-group and community-based studies. *AIDS* 1998;**12**(suppl A):S209-15.

Kelly 1997b

Kelly JA, Murphy DA, Sikkema KJ, et al.Randomised, controlled, community-level HIV-prevention intervention for sexual-risk behavior among homosexual men in US cities. *Lancet* 1997;**350**:1500-5.

Kleinbaum 1998

Kleinbaum DG, Kupper LL, Muller KE, et al.*Applied regression analysis and other multivariable methods*. Pacific Grove: Duxbury Press, 1998.

Lancet 2000

Editorial. What if we had an AIDS vaccine?. *Lancet* 2000;**356**:85.

Leone 2004

Leone P, Hightow L, Foust E, Owen-O'Dowd J, Phillip S, Gray P, et al.HIV transmission among black college student and non-student men who have sex with men --- North Carolina, 2003. *MMWR Morbid Mortal Wkly Report* 2004;**53**:731–734.

Lister 2003

Lister NA, Smith A, Tabrizi S, Hayes P, Medland NA, Garland S, Fairley CK. Screening for Neisseria gonorrhoeae and Chlamydia trachomatis in men who have sex with men at male-only saunas. *Sexually Transmitted Diseases* 2003;**30**:886–889.

Macaskill 2001

Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine* 2001;**20**: 641–654.

MacKellar 2002

MacKellar DA, Valleroy LA, Secura GM. Repeat HIV testing, risk behaviors, and HIV seroconversion among young men who have sex with men: a call to monitor and improve the practice of prevention. *J Acquir Defic Syndr* 2002;**29**(1):76–85.

MacKellar 2005

MacKellar DA, Valleroy LA, Secura GM, Behel S, Bingham T, Celentano DD, et al.Unrecognized HIV infection, risk behaviors, and perceptions of risk among young men who have sex with men. *J Acquir Immune Defic Syndr* 2005;**38**(5):603–614.

MacKellar 2006

MacKellar DA, Valleroy LA, Behel S, et al.Unintentional HIV exposures from young men who have sex with men who disclose being HIV-negative. *AIDS* 2006;**20**:1637–1644.

MacKellar 2007

MacKellar DA, Valleroy LA, Secura GM, et al.Perceptions of lifetime risk and actual risk for acquiring HIV among young men who have sex with men. *AIDS Behav* 2007;**11**:263–270.

Mao 2006

Mao L, Crawford JM, Hospers HJ, et al.'Serosorting' in casual anal sex of HIV-negative gay men is noteworthy and is increasing in Sydney, Australia. *AIDS* 2006;**20**:1204–1206.

Markel 2005

Markel H. The Search for Effective HIV Vaccines. *N Engl J Med* 2005;**335**(8):753–757.

McAuliffe 1999

McAuliffe TL, Kelly JA, Sikkema KJ, Murphy DA, Winett RA, Solomon LJ, Roffman RA. Sexual HIV risk behavior levels among young and older gay men outside of AIDS epicenters: findings of a 16-city sample. *AIDS and Behavior* 1999;**3**:111–119.

McFarland 1999

McFarland W, Busch MP, Kellogg TA, et al.Detection of early HIV infection and estimation of incidence using a sensitive/less-sensitive enzyme immunoassay testing strategy at anonymous counseling and testing sites in San Francisco. *J Acquir Immune Defic Syndr* 1999;**22**: 484–9.

McIntosh 1996

McIntosh MW. The population risk as an explanatory variable in research synthesis of clinical trials. *Stat Med* 1996;**15**:1713-28.

McLean 2006

McLean C. STD and HIV prevention in men who have sex with men: back to basics. *Sexually Transmitted Diseases* 2006;**33**(8):474–475.

Millett 2004

Millett G. More effort needed with 'down low' males. CDC expert discusses issue in Q & A. *AIDS Alert* 2004;**19**:42–43.

Millett 2007

Millett GA, Peterson JL. The known hidden epidemic: HIV/AIDS among black men who have sex with men in the United States. *Am J Prev Med* 2007;**32**(4S):S31–S33.

Mills 1997

Mills TC, Stall RD, Catania JA, et al.Interpreting HIV prevalence and incidence among Americans: bridging data and public policy. *Am J Public Health* 1997, (87):864-6.

Mullen 2002

Mullen PD, Ramirez G, Strouse D, et al.Meta-analysis of the effects of behavioral HIV prevention interventions on the sexual risk behavior of sexually experienced adolescents in controlled studies in the United States. *J Acquir Immune Defic Syndr* 2002;**30**(suppl 1):S94-105.

Murray 1998

Murray DM. *Design and Analysis of Group-Randomized Trials*. New York: Oxford University Press, 1998.

Neumann 2002

Neumann MS, Johnson WD, Semaan S, et al.Review and meta-analysis of HIV prevention intervention research for heterosexual adult populations in the United States. *J Acquir Immune Defic Syndr* 2002;**30**(suppl 1):S106-117.

NIH 2000

National Institutes of Health, National Institute of Allergy and Infectious Diseases. HIV vaccine development: status report. http://www.niaid.nih.gov/daids/vaccine/ whsummarystatus.htm. May 2000; accessed February 27, 2002.

NIH 2002

National Institutes of Health, Office of AIDS Research. Fiscal year 2003 plan for HIV-related research VI: Behavioral and social science. http://www.nih.gov/od/oar/public/pubs/fy2003/vi_behavioral.pdf. Accessed February 27, 2002.

O'Leary 1997

O'Leary A, DiClemente RJ, Aral SO. Reflections on the design and reporting of STD/HIV behavioral intervention research. *AIDS Educ Prevent* 1997;**9**(suppl A):1-14.

Oakley 1995

Oakley A, Fullerton D, Holland J. Behavioral interventions for HIV/AIDS prevention. *AIDS* 1995;**9**:479-86.

Olkin 1998

Olkin I. Odds ratios revisited. *Evidence-Based Medicine* 1998;**3**:71.

Page-Shafer 2000

Page-Shafer K, Kim A, Norton P, et al.Evaluating national HIV prevention indicators: a case study in San Francisco. *AIDS* 2000;**14**(13):2015–26.

Patterson 2003 drug

Patterson TL, Semple SJ. Sexual risk reduction among HIV-positive drug-using men who have sex with men. *J Urban Health* 2003;**80**(iii):77–87.

Peterson 1992

Peterson JL, Coates TJ, Catania JA, et al.High-risk sexual behavior and condom use among gay and bisexual African-American men. *Am J Public Health* 1992;**82**:1490–1494.

Pinkerton 1997

Pinkerton SD, Holtgrave DR, Valdiserri RO. Cost-effectiveness of HIV-prevention skills training for men who have sex with men. *AIDS* 1997;**11**:347-57.

Pinkerton 1998

Pinkerton SD, Holtgrave DR, DiFranceisco WJ, et al.Cost-effectiveness of a community-level HIV risk reduction intervention. *Am J Public Health* 1998;**88**:1239-42.

Pinkerton 2001

Pinkerton SD, Johnson-Masotti AP, Holtgrave DR, et al.Using cost-effectiveness league tables to compare interventions to prevent sexual transmission of HIV. *AIDS* 2001;**15**:917–28.

Ramirez-Valles 2007

Ramirez-Valles J. The quest for effective HIV-prevention interventions for Latino gay men. *Am J Prev Med* 2007;**32**(4S):S34–S35.

Read 2007

Read TRH, Hocking J, SInnott V, Hellard M. Risk factors for incident HIV infection in men having sex with men: a case control study. *Sexual Health* 2007;**4**:35–39.

Rebchook 2006

Rebchook GM, Kegeles SM, Huebner DM, and the TRIP Research Team. Translating research into practice: the dissemination and initial implementation of an evidence-based HIV prevention program. *AIDS Educ Prevent* 2006;**28**(Supplement A):119–136.

Remis 2000

Remis R, Alary M, Otis J, et al.HIV incidence and correlates among homosexual men in the Montreal OMEGA cohort study: The role of unprotected anal and oral sex. XIII International AIDS Conference. Durban, South Africa, 2000. [: abstract MoPpC1100]

Renzi 2003

Renzi C, Tabet SR, Stucky JA, et al.Safety and acceptability of the Reality condom for anal sex among men who have sex with men. *AIDS* 2003;**17**(5):727–31.

Richens 2000

Richens J, Imrie J, Copas A. Condoms and seat belts: the parallels and the lessons. *Lancet* 2000;**355**:400-3.

Rietmeijer 2007

Rietmeijer CA, Lloyd LV, McLean C. Discussing HIV Serostatus With Prospective Sex Partners: A Potential HIV Prevention Strategy Among High-Risk Men Who Have Sex With Men. *Sexually Transmitted Diseases* 2007;**34**(4):215–219.

Robinson 1999

Robinson EK, Evans BG. Oral sex and HIV transmission. *AIDS* 1999;**13**:737-8.

Rosenthal 1994

Rosenthal R. Parametric measures of effect size. In: Cooper H, Hedges LV editor(s). *The handbook of research synthesis*. New York: Russell Sage Foundation, 1994:231–244.

Rummel 1970

Rummel RJ. *Applied factor analysis*. Evanston: Northwestern University Press, 1970.

Sackett 1996

Sackett DL, Deeks JJ, Altman D. Down with odds ratios!. *Evidence-Based Medicine* 1996;**1**:164–167.

Sanchez 2000

Sanchez J, Ojeda J, Garcia P, et al.Incidence and risk factors for HIV acquisition among men who have sex with other men: the Alaska cohort of Lima. XIII International AIDS Conference. Durban, South Africa, 2000. [: abstract TuPeC3380]

Sanchez 2007

Sanchez J, Lama JR, Kusunoki L, et al.HIV-1, sexually transmitted infections, and sexual behavior trends among men who have sex with men in Lima, Peru. *J Acquir Immune Defic Syndr* 2007;**44**:578–585.

Sanchez-Meca 2003

Sanchez-Meca J, Marin-Martinez F, Chacon-Moscoso S. Effect-size indices for dichotomized outcomes in meta-analysis. *Psychol Methods* 2003;**8**:448–467.

Schmid 1998

Schmid CH, Lau J, McIntosh MW, et al.An empirical study of the effect of the control rate as a predictor of treatment efficacy in meta-analysis of clinical trials. *Stat Med* 1998;**17**:1923-42.

Semaan 2002a

Semaan S, Des Jarlais DC, Sogolow E, et al.A meta-analysis of the effect of HIV prevention interventions on the sex behaviors of drug users in the United States. *J Acquir Immune Defic Syndr* 2002;**30**(suppl 1):S73-93.

Semaan 2002b

Semaan S, Kay L, Strouse D, et al.A profile of U.S.-based trials of behavioral and social interventions for HIV risk reduction. *J Acquir Immune Defic Syndr* 2002;**30**(suppl 1):S30-50.

Semple 2003

Semple SJ, Patterson TL, Grant I. Binge use of methamphetamine among HIV-positive men who have sex with men: pilot data and HIV prevention implications. *AIDS Education & Prevention* 2003;**15**:133–47.

Senn 1998

Senn S. Odds ratios revisited. *Evidence-Based Medicine* 1998;**3**:71.

Senterfitt 1998

Senterfitt JW, Harawa NT, Whitfeld C, et al.Study of acute and early infections yields surprising risk profile. 5th Conference on Retrovirology and Opportunistic Infection. Chicago, 1998. [: abstract no. 141]

Slavin 2007

Slavin S, Batrouney C, Murphy D. Fear appeals and treatment side-effects: An effective combination for HIV prevention?. *AIDS Care* 2007;**19**(1):130–137.

Sogolow 2002

Sogolow E, Peersman G, Semaan S, et al.The HIV/AIDS Prevention Research Synthesis project: scope, methods and study classification results. *J Acquir Immune Defic Syndr* 2002;**30**(suppl 1):S15–29.

Solomon 1986

Solomon MZ, DeJong W. Recent sexually transmitted disease prevention efforts and their implications for AIDS health education. *Health Education Quarterly* 1986;**13**(4):301–16.

Stall 1990

Stall R, Ekstrand M, Pollack L, et al.Relapse from safer sex: the next challenge for AIDS prevention efforts. *J Acquir Immune Defic Syndr Hum Retrovirol* 1990;**3**:1181-7.

Stekler 2006

Stekler J, Collier AC, Holmes KK, Golden MR. Primary HIV infection education: knowledge and attitudes of HIV-negative men who have sex with men attending a public health sexually transmitted disease clinic. *J Acquir Immune Defic Syndr* 2006;**42**:123–126.

Stephenson 2000

Stephenson JM, Imrie J, Sutton SR. Rigorous trials of sexual behavior interventions in STD/HIV prevention: what can we learn from them?. *AIDS* 2000;**14**(suppl 3):S115-24.

Stott 1999

Stott J, Hahn BH. Vaccines and immunology. *AIDS* 1999;**13**(suppl A):S103-4.

Sweeney 1995

Sweeney P, Lindegren ML, Buehler JW, Onorato IM, Janssen RS. Teenagers at risk of HIV infection: Results from seroprevalence surveys in the United States. *Archives of Pediatrics & Adolescent Medicine* 1995;**149**(5):521–8.

Tao 1998

Tao G, Remafedi G. Economic evaluation of an HIV prevention intervention for gay and bisexual male adolescents. *J Acquir Immune Defic Syndr Hum Retrovirol* 1998;**17**:83-90.

Torian 2001

Torian LV, Koblin BA, Guilin VA, et al.High HIV seroprevalence and race differentials in young men who have sex with men, sampled at public venues in New York City, 1998-1999. The 8th Conference on Retroviruses and Opportunistic Infections. Chicago, 2001. [: abstract 212]

UNAIDS.org

www.UNAIDS.org website. 2002; accessed February 27, 2002. [: www.UNAIDS.org]

Valleroy 2000

Valleroy LA, MacKellar DA, Karon JM, et al.HIV prevalence and associated risks in young men who have sex with men. *JAMA* 2000;**284**(2):198–204.

Vittinghoff 1999

Vittinghoff E, Douglas J, Judson F, et al.Per-contact risk of human immunodeficiency virus transmission between male sexual partners. *Am J Epidemiol* 1999;**150**:306-11.

Waldo 2000

Waldo CR, McFarland W, Katz MH, et al.Very young gay and bisexual men are at risk for HIV infection: the San Francisco Bay Area Young Men's Survey II. *J Acquir Immune Defic Syndr* 2000;**24**:168–74.

Walter 1997

Walter SD. Variation in baseline risk as an explanation of heterogeneity in meta-analysis. *Stat Med* 1997;**16**:2883-900.

Wawer 2005

Wawer MJ, Gray RN, Sewankambo NK, et al.Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *J Infect Dis* 2005;**191**:1403–1409.

Weinstock 1996

Weinstock HS, Sidhu J, Gwinn M, et al.Trends in HIV seroprevalence among persons attending sexually transmitted disease clinics in the United States, 1988-1992. *JAIDS* 1995;**9**(5):514–22.

Wolitski 2001

Wolitski RJ, Valdiserri RO, Denning PH, et al.Are we headed for a resurgence in the HIV epidemic among men who have sex with men?. *Am J Public Health* 2001;**91**:883-8.

Wolitski 2006

Wolitski RJ, Jones KT, Wasserman JL, Smith JC. Self-identification as "down low" among men who have sex with men (MSM) from 12 US cities. *AIDS Behav* 2006;**10**:519–529.

Woods 2007

Woods WJ, Binson D, Blair J, Han L, Spielberg F, Pollack LM. Probability sample estimates of bathhouse sexual risk behavior. *J Acquir Immune Defic Syndr* 2007;**45**:231–238.

Xia 2006

Xia Q, Molitor F, Osmond DH, THolandi M, Pollack LM, Ruiz JD, Catania JA. Knowledge of sexual partner's HIV serostatus and serosorting practices in a California population-based sample of men who have sex with men. *AIDS* 2006;**20**:2081–2089.

Zea 2003

Zea MC, Reisen CA, Diaz RM. Methodological issues in research on sexual behavior with Latino gay and bisexual men. *American Journal of Community Psychology* 2003;**31**:281–291.

Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men (Review)   **34**

Copyright © 2009 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Zenilman 2005

Zenilman JM. Behavioral interventions - rationale, measurement, and effectiveness. *Infect Dis Clin N Am* 2005;**19**:541–562.

References to other published versions of this review

Johnson 2002a

Johnson WD, Hedges LV, Ramirez G, et al.HIV prevention research for men who have sex with men: A systematic review and meta-analysis. *J Acquir Immune Defic Syndr* 2002;**30**(Suppl 1):S118–129.

Johnson 2003

Johnson WD, Hedges LV, Diaz RM. Interventions to modify sexual risk behaviors for preventing HIV infection in men who have sex with men. *Cochrane Database of Systematic Reviews* 2003;**1**:CD001230.

Johnson 2005

Johnson WD, Holtgrave DR, McClellan WM, Flanders WD, Hill AN, Goodman M. HIV Intervention Research for Men Who Have Sex with Men: a 7-Year Update. *AIDS Education and Prevention* 2005;**17**(6):568-589.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies   *[ordered by study ID]*

Carballo-Diéguez2005

Methods	random assignment
Participants	141 Latino MSM in New York
Interventions	8 sessions on themes of oppression, transgression of rules, excuses (or rationalizations), substance use, goal setting, the role of pleasure, self-efficacy and plans for the future. Exercises included word association, story analysis, problem solving, analysis of Spanish dichos (proverbs), discussion of participants weekly sexual diaries
Outcomes	Occasions of UAI; any UAI in past 2 mo
Notes	wait list control

CDC ACDP 1999@

Methods	series of cross-sectional surveys in matched communities
Participants	536 MSM who do not self-identify as gay, surveyed in public sex environments in Seattle, Denver, and southern California 1991-95

CDC ACDP 1999@ *(Continued)*

Interventions	AIDS Community Demonstration Projects: Community level intervention in Seattle and East Denver featuring distribution and discussion of flyers containing condoms and role-model stories from men in the community about making progress toward consistent condom use (32 months in community)
Outcomes	Any UAI in past 1 mo
Notes	No treatment in the paired control communities (Long Beach and West Denver)

Choi 1996@

Methods	random assignment
Participants	256 self-identified homosexual Asian or Pacific Islander men in San Francisco, 1992-94
Interventions	API Living Well Project: Culturally specific brief group counseling including development of positive self-identity and social support, safer sex education, eroticizing and negotiating safer sex. One 3-hr session.
Outcomes	Any UAI in past 3 mo
Notes	wait list control

Cleary 1995*@

Methods	random assignment
Participants	112 MSM blood donors testing HIV-positive
Interventions	Individual counseling (IC) plus a cognitive behavioral and skills training support group to provide more detailed info, encourage risk-reduction behavior, provide support and facilitate functional coping responses. 6 weekly meetings of 90 minutes
Outcomes	Any UAI in past 2 weeks
Notes	IC plus community referral

Coates 1989@

Methods	random assignment
Participants	64 asymptomatic HIV-seropositive MSM who were not already practicing meditation regularly. San Francisco, 1987.
Interventions	stress mgt skills, systematic relaxation, health behavior change (8 sessions of 2 hrs and 1 all-day retreat)

Coates 1989@ (*Continued*)

Outcomes	Any UAI in past 1 mo
Notes	wait list control

Dilley 2002 d*@

Methods	random assignment to 1 of 4 conditions; all kept diary of sexual activity
Participants	138 MSM, San Francisco, 1997-2000
Interventions	Individual standard counseling (ISC, one 1-hr session) plus self-justifications (SJ) session, where the client reviewed and challenged his own self-justifications for a recent occasion of unsafe sex, AND diary of sexual activity for 90 days (labeled B2 in article)
Outcomes	Any UAI in past 3 mo
Notes	ISC + diary but no SJ (labeled A2 in article)

Dilley 2002 n*@

Methods	same study as above without diary
Participants	
Interventions	ISC (one 1-hr session) plus SJ session, with no diary (labeled B1 in article)
Outcomes	
Notes	ISC only, no SJ or diary (labeled A1 in article)

Dilley 2007*#

Methods	random assignment
Participants	305 MSM attending San Francisco HIV CT clinic, 2002-04
Interventions	Individual personalized cognitive counseling by a paraprofessional along with usual CT
Outcomes	Occasions of UAI with 2 most recent potentially serodiscordant noncommitted partners in past 3 mo
Notes	Control received usual CT only

Elford 2001@

Methods	staggered implementation across 5 gyms; serial cross-sectional survey
Participants	1010 MSM in 5 gyms in London, 1997-98
Interventions	Community level intervention: replication of Kelly 1991 in 5 gyms. Difficulties reported in delivering the intervention; only 3% of respondents reported having been spoken to by volunteers.
Outcomes	Any UAI in past 3 mo
Notes	lagged design (for control gyms)

Explore 2004*

Methods	random assignment
Participants	3775 MSM in 6 US cities 1999-2003
Interventions	Ten 1-on-1 counseling sessions followed by maintenance sessions every 3 mo. Risk assessment, sexual communication, knowledge of HIV serostatus, alcohol and drug use, triggers for unsafe sex, motivational interviewing. Total span up to 48 mo
Outcomes	Occasions of UAI; any UAI in past 6 mo
Notes	Control condition was twice-yearly counseling & HIV testing

Flowers 2002

Methods	Glasgow assigned to intervention. Edinburgh to control. Behaviors measured in 2 cross sectional surveys
Participants	2271 men at gay bars in Glasgow, Edinburgh 1996-99
Interventions	Gay Men's Task Force: Community-level intervention delivered through gay bars in Glasgow. Peer-led sex health promotion, gay-specific genitourinary medicine services, free phone hotline w/sex health info & details of local sexual health services, endorsement of testing, risk assessment, and sexual health. 9 months in community, 2 days training for peer educators
Outcomes	Occasions of UAI; any UAI in past 12 mo
Notes	No treatment in Edinburgh

Gold 1995 po*

Methods	random assignment to 1 of 3 conditions
Participants	109 gay men who had recently had UAI. Melbourne and Sydney, 1993
Interventions	Individual level. Diary of sexual behavior for 16 weeks, plus (at 4 weeks) examination of posters used in AIDS education (10 posters)
Outcomes	Occasions of UAI; any UAI in past 12 weeks
Notes	16-week sexual diary only

Gold 1995 sj*

Methods	same study as above
Participants	
Interventions	Individual level. Diary of sexual behavior for 16 weeks, plus (at 4 weeks) evaluation of their own self- justifications for an occasion of UAI (2 exercises)
Outcomes	
Notes	16-week sexual diary only

Gold 1998 rse*

Methods	random assignment to 1 of 3 conditions; all were assigned to keep a diary
Participants	92 MSM who had recently had UAI; Melbourne and Sydney, 1996
Interventions	Individual level. Assigned to describe in detail a recent sexual encounter that included anal sex without a condom
Outcomes	Occasions of UAI; any UAI in past 12 weeks
Notes	Diary only

Gold 1998 sjp*

Methods	same study as above
Participants	

Gold 1998 sjp* *(Continued)*

Interventions	Individual level. Assigned to examine and describe 10 AIDS education posters highlighting the pitfalls of SJ
Outcomes	
Notes	Diary only

Harding 2004

Methods	random assignment
Participants	19 MSM in London 2000
Interventions	'SM sex: an introduction to the SM scene'. Sessions address assumptions and knowledge, practical tools of SM sex, risk taking, emotional aspects, sexually transmitted infections and HIV transmission, rights and responsibilities, legal issues, the role of fantasy, and limits and boundaries. Up to 25 group members, 4 sessions of 7 hrs
Outcomes	Occasions of UAI; any UAI in past 3 mo
Notes	Wait list

Healthy Living 2007#

Methods	random assignment
Participants	936 HIV+ people in Los Angeles, Milwaukee, New York, and San Francisco. 57% were MSM
Interventions	Individual level. 15 90-minute sessions in 3 modules: stress, coping, adjustment; safer behaviors; and health behaviors
Outcomes	Occasions of UAI in past 3 mo at 5, 10, 15, 20, and 25 mo after randomization
Notes	Wait list

Hoff 1997

Methods	Portland assigned to intervention, Tucson to control
Participants	537 MSM in Portland, OR and Tucson, AZ 1992-1996
Interventions	The Portland intervention targeted community mobilization, social support, education, outreach, volunteer coordination, HIV testing, and provider mobilization. 18 months in community
Outcomes	Occasions of UAI; any UAI in past 1 mo, past 12 mo

Hoff 1997 *(Continued)*

Notes	No treatment in Tucson

Imrie 2001*@

Methods	random assignment
Participants	252 gay men attending a sexual health clinic with acute STI or unprotected sex in past year. London 1995-98
Interventions	Gay Men Project: standard mgt (1-to-1 counseling & referrals, 20 minutes) plus 1-day small group workshop
Outcomes	Any UAI in past 1 mo, past 12 mo
Notes	Standard management only

Kalichman 2001#

Methods	random assignment
Participants	164 MSM with HIV (62% of participants were MSM, 74% African Americans), Atlanta 1997
Interventions	Support group to create sexual health and relationship plans, develop communication and disclosure skills, learn hazards of co-infection with other STI. Five 120-min sessions
Outcomes	Occasions of UAI in past 3 mo
Notes	Support group for health maintenance. Five 120-min sessions

Kegeles 1996@

Methods	Eugene (Oregon) assigned to intervention, Santa Barbara (California) to control
Participants	Cohort of 100 young gay men in intervention community (Eugene, Oregon) and 88 in comparison community (Santa Barbara, California) 1993
Interventions	Mpowerment Project: Community-level peer-led program for young gay men including outreach, small groups, and a publicity campaign (8 months in community)
Outcomes	Any UAI in past 2 mo
Notes	lagged design (for comparison community)

Kegeles 2002@

Methods	Albuquerque assigned to treatment, Austin and Phoenix to control
Participants	632 young MSM in Albuquerque, Austin, Phoenix 1997-98
Interventions	Based on theories of empowerment, diffusion and peer mobilization, the Albuquerque intervention featured a young gay men's community center, a core group of men who ran the project, informal outreach among friends, formal outreach at gay venues and social events, and small groups focused on safer sex and informal outreach. 12 months in community
Outcomes	Any UAI in past 2 mo
Notes	No treatment in Austin and Phoenix

Kelly 1989#

Methods	random assignment of individuals
Participants	85 MSM in Jackson, Mississippi, 1987
Interventions	Project ARIES: AIDS risk education, cognitive-behavioral self-mgt training, sexual assertion training, development of relationship skills, and social support (12 weekly meetings of 75-90 min each)
Outcomes	Occasions of UAI in past 4 mo
Notes	wait list

Kelly 1991

Methods	randomized lagged design of intervention delivery to communities; serial cross-sectional survey
Participants	634 MSM at gay bars in 3 communities in Mississippi and Louisiana, 1989
Interventions	Popular Opinion Leader (POL): Community level intervention in which popular opinion leaders were trained to endorse behavior change to peers in gay clubs (4 wkly training sessions, 90 minutes each)
Outcomes	Occasions of UAI; any UAI in past 2 mo
Notes	lagged design (for 2 control communities)

Kelly 1993 cb

Methods	random assignment to 1 of 3 conditions
Participants	69 depressed HIV-positive men in Milwaukee 1991
Interventions	Milwaukee AIDS Project: Cognitive behavioral approach with behavioral or skill training themes (8 wkly 90-minute sessions)
Outcomes	Occasions of UAI; any UAI in past 3 mo
Notes	

Kelly 1993 ss

Methods	same study as above
Participants	
Interventions	Social support group (8 wkly 90-minute sessions)
Outcomes	
Notes	Crisis therapy only if requested

Kelly 1997

Methods	random assignment to intervention within each of 4 pairs of communities; serial cross sectional survey
Participants	386 MSM in 4 pairs of communities in Wisconsin, New York, West Virginia, and Washington 1991-94
Interventions	Scale-up of POL. Community level intervention in gay bars. Popular men were engaged to advocate benefits of behavior change to peers, and HIV education materials were placed in bars (5 wkly training sessions of 2 hrs each for opinion leaders)
Outcomes	Occasions of UAI; any UAI in past 2 mo
Notes	Educational materials in bars only

Miller 1998

Methods	lagged design across 3 neighborhood bars; serial cross-sectional survey
Participants	385 men (57% identified as gay, 31% as bisexual) men at hustler bars in New York City, 1996

Miller 1998 (*Continued*)

Interventions	Hustler Bar Project: Community level intervention. Replication of Kelly 1991 in hustler bars
Outcomes	Occasions of UAI; any UAI in past 2 mo
Notes	lagged design (2 control bars)

Patterson 2003 b#

Methods	random assignment to 1 of 4 conditions
Participants	286 HIV+ volunteers recruited by posters, service providers and others, and reporting unprotected sex with HIV-negative or unknown status partners. These 286 who identified as gay or bisexual constituted 85% of the followup respondents. San Diego 1999-2001
Interventions	Booster-enhanced social cognitive intervention in 3 domains (condom use, negotiation of safer sex, disclosure of HIV status). One 90-min comprehensive session plus two 90-min booster sessions
Outcomes	Occasions of UAI in past 4 mo
Notes	Three 90-min sessions on diet and exercise

Patterson 2003 c#

Methods	same study as above
Participants	
Interventions	Comprehensive social cognitive intervention in 3 domains (condom use, negotiation of safer sex, disclosure of HIV status). 90 minutes
Outcomes	
Notes	Three 90-min sessions on diet and exercise

Patterson 2003 s#

Methods	same study as above
Participants	
Interventions	Targeted social cognitive intervention in 1 of 3 domains (condom use, negotiation of safer sex, disclosure of HIV status) selected by participant. 90 minutes

Patterson 2003 s# *(Continued)*

Outcomes	
Notes	Three 90-min sessions on diet and exercise

Peterson 1996 1s

Methods	random assignment of small groups of consecutively enrolled individuals to 1 of 3 conditions
Participants	177 African American homosexual and bisexual men in San Francisco and Oakland 1989-91
Interventions	1 session on AIDS risk education, cognitive- behavioral self- management training, assertion training, self-identity and social support (one 3-hr session)
Outcomes	Occasions of UAI; any UAI in past 6 mo
Notes	wait list

Peterson 1996 3s

Methods	same study as above
Participants	
Interventions	3 sessions on AIDS risk education, cognitive- behavioral self- management training, assertion training, self-identity and social support (3 weekly 3-hr sessions)
Outcomes	
Notes	wait list

Picciano 2001#

Methods	random assignment
Participants	89 MSM who reported 3 or more recent episodes of oral or anal sex without condoms; Seattle 1998-99
Interventions	Feedback by telephone regarding a baseline risk assessment (1 hour)
Outcomes	Occasions of UAI in past 6 wk
Notes	wait list

R-B 2004 in person@

Methods	random assignment to 1 of 3 conditions
Participants	121 substance-using young MSM with HIV, most referred by social service agencies or medical providers. The 121 MSM constituted 69% of the followup respondents. Los Angeles, New York, San Francisco, 1999-2002
Interventions	Individual format modification of RB 2001 to increase participation. Improving physical health, maintaining drug regimens, coping with learning HIV status, health care decisions. Reducing unprotected sex and substance use, examining trigger situations, condom use and negotiation skills and self-efficacy. Focus on condom use rather than disclosure. Reducing distress, anticipating situations that raise anxiety, depression, fear, or anger. Recognizing and controlling negative emotion with relaxation, self-instruction, meditation. Identifying life goals. 18 weekly 2-hr sessions. Delivered in person
Outcomes	Any UAI in past 3 mo
Notes	wait list

R-B 2004 phone@

Methods	same study as above
Participants	
Interventions	Same as above, but delivered by telephone
Outcomes	
Notes	wait list

Read 2006*#

Methods	random assignment
Participants	110 MSM age 18+ who receive HIV negative test results at the Hollywood gay service center [year?]
Interventions	Individual level. Interactive video (IAV) with peer counseling vs peer counseling alone. IAV designed to simulate the emotional, interpersonal, and contextual narrative of a real sexual encounter while challenging and changing risky responses
Outcomes	Occasions of UAI past 3 mo
Notes	Peer counseling only

Richardson 2004 g@

Methods	2 clinics were assigned to each of 3 conditions gain frame (G), loss frame (L) or control
Participants	402 MSM patients at 6 HIV treatment clinics, California 1999
Interventions	Two clinics assigned to use a gain-framed approach (G) (positive consequences of safer-sex). Prevention counseling from medical providers supplemented with written information. 3-5 min every visit; 4 hrs training for clinicians
Outcomes	Any UAI in past 3 mo
Notes	2 attention- control clinics were assigned to medication adherence intervention

Richardson 2004 L@

Methods	same study as above
Participants	
Interventions	Same as above, but two clinics were assigned to a loss-frame (L) approach (negative consequences of unsafe sex)
Outcomes	
Notes	same as above

Roffman 1997

Methods	random assignment
Participants	410 MSM in western US, 1992-94
Interventions	Project ARIES: Small group telephone conferences of geographically dispersed MSM; permitted anonymity and participation by men in rural areas. Coping strategies to deal with high-risk situations, setting realistic, client-centered risk reduction goals, identifying antecedents to risk behavior. 14 weekly 90-min phone calls
Outcomes	Occasions of UAI; any UAI in past 4 wk
Notes	wait list

Roffman 1998#

Methods	alternating assignment
Participants	129 MSM at risk of relapse to unsafe sex, Seattle 1989-91

Roffman 1998# *(Continued)*

Interventions	Relapse prevention: HIV education, motivation, listening, self-talk, assertiveness, avoiding risky situations, debriefing, maintenance strategies, social support, self-esteem. 17 weekly sessions of 2 hours each
Outcomes	Occasions of UAI in past 3 mo
Notes	wait list

Rosser 1990 c@

Methods	random assignment
Participants	139 sexually active gay men in Auckland, New Zealand, 1987-88
Interventions	Individual HIV prevention counseling using a behavioral HIV risk assessment system (30 min)
Outcomes	Any UAI past 2 mo outside a mutually monogamous relationship
Notes	no treatment

Rosser 1990 e@

Methods	same study as above
Participants	
Interventions	Small group workshop on eroticizing safer sex (2 to 2.5 hrs)
Outcomes	
Notes	no treatment

Rosser 1990 s@

Methods	same study as above
Participants	
Interventions	Small group Stop AIDS workshop (2 to 2.5 hrs)
Outcomes	
Notes	no treatment

Rosser 1990 v@

Methods	same study as above
Participants	
Interventions	video on safer sex (15 min)
Outcomes	
Notes	no treatment

Rosser 2002*@

Methods	random assignment
Participants	169 MSM in Minneapolis 1997-98
Interventions	Minnesota Men's Study: comprehensive seminar featuring systematic desensitization, study of homosexual identity formation, sexual health education, research on cofactors of unsafe sex (drugs, alcohol, loneliness, falling in love). 2 full-day sessions
Outcomes	Any UAI in past 3 mo
Notes	3-hour HIV prevention video

Rotheram-Borus 2001@

Methods	
Participants	94 young MSM with HIV at 4 clinical care sites. (63% of the total sample size). Los Angeles, New York, San Francisco, Miami 1994-1997
Interventions	First module Staying Healthy, e.g., coping with learning HIV status, disclosure, health care decisions. Second module Acting Safe, e.g., protecting self and partner, safer sex options, drugs and alcohol, avoiding internal and external triggers, anxiety and anger. Total of 23 weekly 2-hr sessions
Outcomes	Any UAI in past 3 mo
Notes	Usual clinical activities

Sampaio 2002*@

Methods	random assignment
Participants	227 MSM in Salvador, Brazil 1998-99
Interventions	Projeto Contato: Safer sex workshop with games, role playing, small group discussion using verbal and nonverbal communication. Basic info, clarification of misconceptions, recognition of risk. Nongenital practices, safe sex in committed vs. other relationships, mechanics of using condoms, strategies for refusing unsafe sex, negotiating new sexual patterns. One 3-4 hr session with 15-20 subjects
Outcomes	Any UAI in past 1 mo
Notes	1-hr lecture & discussion on infectious disease, condom skills

Shepherd 1997

Methods	one community (Southampton) chosen for intervention, another for control; cohorts enrolled and surveyed in each
Participants	54 young MSM (mean age 24 years) in gay-friendly environments, Southampton (England) 1996
Interventions	Community level. The HAPEER Project: Young peer educators administered a structured interview to peers to recruit for study, collect baseline data, initiate discussion of sexual health, and identify and respond to individual sexual health needs
Outcomes	Occasions of UAI; any UAI in past 3 mo
Notes	Comparison group: young MSM in a neighboring gay community

Shoptaw 2005 cbt*

Methods	random assignment to 1 of 4 conditions
Participants	162 meth-dependent MSM in Los Angeles, 1998-2000
Interventions	CBT+CR: both treatments simultaneously
Outcomes	Occasions of UAI; any UAI in past 1 mo
Notes	CR: Vouchers (e.g., for groceries, camera equipment, plane fare to visit family, clothing) for drug abstinence. Escalating value for successive negative urine samples with reset after relapse; max total value $1300

Shoptaw 2005 g*

Methods	same study as above
Participants	
Interventions	Gay-specific enhancement of CBT. Difference in sexual behavior on and off drug, indicators of meth use in sexual partners and friends, comparison of revealing one's drug problem to the coming out process, examples from circuit parties. 90 min 3 times / wk
Outcomes	
Notes	CBT: group education to initiate meth abstinence and quickly resume abstinence if relapse occurs. Internal and external triggers, stages of recovery from meth dependence, identification of emotional states signaling relapse. Cognitive skills such as thought stopping, craving management, relapse analysis, adoption of healthy lifestyle behaviors. 90 min 3 times / wk

Sorensen 2003*

Methods	random assignment
Participants	42 substance abusers with HIV/AIDS at a teaching and public hospital from inpatient medical wards, outpatient heroin detox clinic, and emergency department. (These 42 MSM constituted 48% of follow-up respondents). San Francisco 1994-1997
Interventions	12 months intensive case management for Hybrid of full service and referral models. Case managers were paraprofessionals, former consumers of HIV or substance abuse treatment services.
Outcomes	Occasions of UAI; any UAI in past 1 mo
Notes	Brief contact: education, risk reduction info, referrals

Stall 1999@

Methods	alternating assignment of small groups
Participants	129 MSM attending a nonresidential treatment center in San Francisco 1990-93
Interventions	Closed group treatment for substance use disorder plus exercises concerning sexual risk-taking (16-week program, two 3-hour sessions per week)
Outcomes	Any UAI past 3 mo
Notes	Group treatment for substance use only

Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men (Review)   **51**

Copyright © 2009 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Tudiver 1992 c4@

Methods	random assignment to 1 of 3 conditions
Participants	500 gay men in Toronto, 1990
Interventions	Talking Sex Project: 4 sessions led by paid counselors. Discussion of safer sex, personal experiences, coping strategies, skills, and role plays (4 weekly 2-hr sessions)
Outcomes	Any UAI in past 3 mo
Notes	wait list

Tudiver 1992 p1@

Methods	same study as above
Participants	
Interventions	1 session led by trained volunteers. Discussion of safer sex, personal experiences, coping strategies, skills, and role plays (one 3-hr session)
Outcomes	
Notes	wait list

Valdiserri 1989*

Methods	random assignment of small groups of consecutively enrolled individuals
Participants	432 homosexual and bisexual men in Pittsburgh, 1986-87
Interventions	AIDS Prevention Project: AIDS information and safer sex lecture followed by skills training, discussion and rehearsal of safer sex negotiation (total 140 min)
Outcomes	Number of partners for UAI; any UAI in past 6 mo
Notes	small group lecture only (60-90 min)

Wolitski 2005*

Methods	random assignment
Participants	811 HIV+ MSM, New York and San Francisco, 2000-02

Wolitski 2005*   *(Continued)*

Interventions	Group activities facilitated by HIV+ peers. Building community for HIV+, information, personal responsibility, assumptions and disclosure of serostatus, communication skills, effects of substance use and behavior on immune system, coping with HIV, mental health. Six 3-hr sessions
Outcomes	Occasions of UAI; any UAI in past 3 mo
Notes	90 min forum and lecture

Interventions were delivered in small group format unless specified as community- or individual-level.

* = comparison condition included standard or other HIV prevention intervention

@ = dichotomous data only

= count data only

(studies with no @ or # sign had both dichotomous and count data)

CBT = cognitive-behavioral therapy

CR = contingent reinforcement (sometimes called contingency management)

ISC = individual standard counseling

Meth = methamphetamine

Mgt = management

MSM = men who have sex with men

SJ = self-justifications

UAI = unprotected anal intercourse

For serial cross-sectional surveys (CDC ACDP 1999, Elford 2001, Flowers 2002, Hoff 1997, Kelly 1991, Kelly 1997, Miller 1998), sample sizes shown are half the total number surveyed, because about half represent baseline and about half represent follow-up measurements. For all other studies, sample sizes shown are the number of respondents at follow-up. Number of respondents are averaged across multiple followup times. Participant sample sizes include only MSM participants. When multiple interventions were evaluated against comparison conditions in a single study (Dilley 2002, Gold 1995, Gold 1998, Kelly 1993, Patterson 2003, Peterson 1996, Rotheram-Borus 2004, Richardson 2004, Rosser 1990, Shoptaw 2005, Tudiver 1992), the total follow-up sample size is shown under "Participants" only in the first entry representing that study.

Characteristics of excluded studies   *[ordered by study ID]*

Antoni 2000	HIV prevention not a focus of the intervention
Blake 2001	Comparison between schools according to 4 levels of pre-existing gay-sensitive instruction; no assignment to intervention condition
Bowen 2007	No behavioral outcomes reported - outcomes are knowledge, outcome expectancies, and self-efficacy. Only 1 week followup before control group received the intervention
Card 2001	Targeted adolescents, not specific to MSM. For use in communities, schools, family planning clinics, STD clinics, mental health centers, and drug rehabilitation centers

Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men (Review)   **53**

Copyright © 2009 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chesney 2003	Intervention focused on coping for HIV-positives rather than behavior change
Cote & Pepler 2002	RCT of coping outcomes only for HIV-positive men
Cote & Pepler 2005	RCT of coping outcomes only for HIV-positive men, secondary to Cote 2002
Cruess 2002	Targeted and measured stress reduction, depressive symptoms, and related skills and attitudes rather than sexual behavior
Dilorio 2007	No occurrence of the words gay, homosexual, bisexual, or sex with males
Fisher 2006	Only 15% of intervention participants and 10% of comparison participants had homosexual sex as route of HIV infection
French 2000	No control group. Gay men in London
Heckman 2004	Intervention focused on coping and distress among HIV-positives rather than sexual risk behavior change
Hospers 1999	No independent comparison group. 362 MSM at cruising areas in the Netherlands.
Huebner 2002	pre-intervention to post-intervention change in attitudes only
Jones 2008	No comparison condition. Adaptation of POL intervention with African American MSM in North Carolina
Kelly 1990	No comparison group data
Miller 1995	Not randomized, and only one time point for comparison group. For studies that were not randomized, we required separate baseline and follow-up for both intervention and comparison groups.
Nokes 2003	Case study of 2 sessions with no control group. 1st session included 5 gay men. 2nd session included 2 of the members from the first session plus 3 more gay men and 1 woman
Perry 1991	No behavioral outcomes reported. Mentioned that no HIV infections occurred during the study
Reback 2004	No comparison group
Remafedi 1994	No comparison group
Rosser 1991	Overlap with Rosser 1990
Toro-Alfonso	No independent comparison group
Ziersch 2000	No data were obtained from the comparison site

(Continued)

Zimmerman 1997	No independent comparison group

Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men (Review)   **55**

Copyright © 2009 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

DATA AND ANALYSES

Comparison 1.   Intervention vs minimal to no HIV prevention

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]	40		Rate ratio (Fixed, 95% CI)	0.73 [0.63, 0.85]
1.1 Small group interventions (Rate ratios)	18		Rate ratio (Fixed, 95% CI)	0.70 [0.55, 0.90]
1.2 Individual-level interventions (Rate ratios)	11		Rate ratio (Fixed, 95% CI)	0.80 [0.60, 1.06]
1.3 Community-level interventions (Rate ratios)	11		Rate ratio (Fixed, 95% CI)	0.70 [0.55, 0.91]
2 PR Proportion reporting any unprotected sex [# = RR substituted for PR]	40		Prevalence ratio (Fixed, 95% CI)	0.77 [0.72, 0.83]
2.1 Small group interventions (Prevalence ratios)	18		Prevalence ratio (Fixed, 95% CI)	0.73 [0.64, 0.84]
2.2 Individual-level interventions (Prevalence ratios)	11		Prevalence ratio (Fixed, 95% CI)	0.84 [0.74, 0.97]
2.3 Community-level interventions (Prevalence ratios)	11		Prevalence ratio (Fixed, 95% CI)	0.75 [0.66, 0.84]

Comparison 2.   Experimental vs standard or other HIV prevention

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]	18		Rate ratio (Fixed, 95% CI)	0.83 [0.73, 0.95]
1.1 Small group interventions (Rate Ratios)	8		Rate ratio (Fixed, 95% CI)	0.77 [0.59, 1.01]
1.2 Individual-level interventions (Rate Ratios)	10		Rate ratio (Fixed, 95% CI)	0.86 [0.73, 1.00]
2 PR Proportion reporting any unprotected sex [# = RR substituted for PR]	18		Prevalence ratio (Fixed, 95% CI)	0.93 [0.89, 0.97]
2.1 Small group interventions (Prevalence ratios)	8		Prevalence ratio (Fixed, 95% CI)	0.87 [0.78, 0.97]

| 2.2 Individual-level interventions (Prevalence ratios) | 10 | Prevalence ratio (Fixed, 95% CI) | 0.94 [0.90, 0.99] |

Analysis 1.1. Comparison 1 Intervention vs minimal to no HIV prevention, Outcome 1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR].

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 1 Intervention vs minimal to no HIV prevention

Outcome: 1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]

Study or subgroup	log [Rate ratio] (SE)	Rate ratio IV,Fixed,95% CI	Weight	Rate ratio IV,Fixed,95% CI
1 Small group interventions (Rate ratios)				
Kelly 1989#	-0.9396 (0.574)		1.8 %	0.39 [0.13, 1.20]
Coates 1989@	0 (0.8196)		0.9 %	1.00 [0.20, 4.98]
Rosser 1990 e@	0.2719 (1.3478)		0.3 %	1.31 [0.09, 18.42]
Rosser 1990 s@	0.7828 (1.353)		0.3 %	2.19 [0.15, 31.02]
Tudiver 1992 p1@	-0.3642 (0.3741)		4.2 %	0.69 [0.33, 1.45]
Tudiver 1992 c4@	-0.0075 (0.4339)		3.1 %	0.99 [0.42, 2.32]
Roffman 1998#	-0.3626 (0.4208)		3.3 %	0.70 [0.31, 1.59]
Peterson 1996 3s	-2.386 (1.1526)		0.4 %	0.09 [0.01, 0.88]
Peterson 1996 1s	-1.8787 (1.1058)		0.5 %	0.15 [0.02, 1.33]
Kelly 1993 cb	0.2574 (0.9153)		0.7 %	1.29 [0.22, 7.78]
Kelly 1993 ss	-1.3571 (1.0345)		0.5 %	0.26 [0.03, 1.96]
Stall 1999@	0.2811 (0.5253)		2.1 %	1.32 [0.47, 3.71]
Roffman 1997	-0.4002 (0.2562)		8.9 %	0.67 [0.41, 1.11]
Choi 1996@	-0.2267 (0.6951)		1.2 %	0.80 [0.20, 3.11]
Rotheram-Borus 2001@	-0.2113 (0.8457)		0.8 %	0.81 [0.15, 4.25]
Kalichman 2001#	-1.1123 (0.5605)		1.9 %	0.33 [0.11, 0.99]
Harding 2004	-0.5509 (1.2099)		0.4 %	0.58 [0.05, 6.17]
Carballo-Diguez2005	-0.1466 (0.3349)		5.2 %	0.86 [0.45, 1.66]
Subtotal (95% CI)			**36.6 %**	**0.70 [0.55, 0.90]**

Heterogeneity: Chi² = 12.98, df = 17 (P = 0.74); I² =0.0%

Test for overall effect: Z = 2.82 (P = 0.0048)

0.2   0.5   1.0   2.0   5.0

Favours treatment   Favours control

(*Continued* . . .)

Study or subgroup	log [Rate ratio] (SE)	Rate ratio IV,Fixed,95% CI	Weight	Rate ratio IV,Fixed,95% CI
2 Individual-level interventions (Rate ratios)				
Rosser 1990 v@	0.5596 (1.3428)		0.3 %	1.75 [0.13, 24.32]
Rosser 1990 c@	-0.4212 (1.3974)		0.3 %	0.66 [0.04, 10.15]
Picciano 2001#	-0.7084 (0.5547)		1.9 %	0.49 [0.17, 1.46]
Patterson 2003 b#	0.0252 (0.4937)		2.4 %	1.03 [0.39, 2.70]
Patterson 2003 c#	-0.2421 (0.4926)		2.4 %	0.78 [0.30, 2.06]
Patterson 2003 s#	0.06 (0.4774)		2.6 %	1.06 [0.42, 2.71]
Richardson 2004 g@	0.0079 (0.4651)		2.7 %	1.01 [0.41, 2.51]
Richardson 2004 L@	-0.3334 (0.4875)		2.5 %	0.72 [0.28, 1.86]
R-B 2004 in person@	0.2519 (1.1209)		0.5 %	1.29 [0.14, 11.57]
R-B 2004 phone@	0.4262 (1.2284)		0.4 %	1.53 [0.14, 17.01]
Healthy Living 2007#	-0.3354 (0.2201)		12.1 %	0.72 [0.46, 1.10]
Subtotal (95% CI)			**28.0 %**	**0.80 [0.60, 1.06]**
Heterogeneity: Chi² = 2.75, df = 10 (P = 0.99); I² =0.0%				
Test for overall effect: Z = 1.54 (P = 0.12)				
3 Community-level interventions (Rate ratios)				
Kelly 1991	-0.1137 (0.3816)		4.0 %	0.89 [0.42, 1.89]
Kelly 1997	-1.3736 (0.343)		5.0 %	0.25 [0.13, 0.50]
Kegeles 1996@	-0.343 (1.0907)		0.5 %	0.71 [0.08, 6.02]
CDC ACDP 1999@	-0.1919 (0.3692)		4.3 %	0.83 [0.40, 1.70]
Hoff 1997	-0.1025 (0.317)		5.8 %	0.90 [0.48, 1.68]
Shepherd 1997	-0.0572 (0.8487)		0.8 %	0.94 [0.18, 4.98]
Miller 1998	-0.4789 (0.4172)		3.4 %	0.62 [0.27, 1.40]
Elford 2001@	0.1169 (0.3676)		4.3 %	1.12 [0.55, 2.31]
Flowers 2002	-0.3185 (0.2998)		6.5 %	0.73 [0.40, 1.31]
Kegeles 2002@	-0.413 (0.9601)		0.6 %	0.66 [0.10, 4.34]
Amirkhanian 2005@	-0.4654 (1.6351)		0.2 %	0.63 [0.03, 15.48]
Subtotal (95% CI)			**35.4 %**	**0.70 [0.55, 0.91]**
Heterogeneity: Chi² = 11.93, df = 10 (P = 0.29); I² =16%				
Test for overall effect: Z = 2.74 (P = 0.0062)				
Total (95% CI)			**100.0 %**	**0.73 [0.63, 0.85]**
Heterogeneity: Chi² = 28.26, df = 39 (P = 0.90); I² =0.0%				
Test for overall effect: Z = 4.15 (P = 0.000033)				
Test for subgroup differences: Chi² = 0.60, df = 2 (P = 0.74), I² =0.0%				

0.2   0.5   1.0   2.0   5.0

Favours treatment   Favours control

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 1 Intervention vs minimal to no HIV prevention

Outcome: 1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]

Study or subgroup	log [Rate ratio] (SE)	Rate ratio IV,Fixed,95% CI	Weight	Rate ratio IV,Fixed,95% CI
1 Small group interventions (Rate ratios)				
Kelly 1989#	-0.9396 (0.574)		1.8 %	0.39 [0.13, 1.20]
Coates 1989@	0 (0.8196)		0.9 %	1.00 [0.20, 4.98]
Rosser 1990 e@	0.2719 (1.3478)		0.3 %	1.31 [0.09, 18.42]
Rosser 1990 s@	0.7828 (1.353)		0.3 %	2.19 [0.15, 31.02]
Tudiver 1992 p1@	-0.3642 (0.3741)		4.2 %	0.69 [0.33, 1.45]
Tudiver 1992 c4@	-0.0075 (0.4339)		3.1 %	0.99 [0.42, 2.32]
Roffman 1998#	-0.3626 (0.4208)		3.3 %	0.70 [0.31, 1.59]
Peterson 1996 3s	-2.386 (1.1526)		0.4 %	0.09 [0.01, 0.88]
Peterson 1996 1s	-1.8787 (1.1058)		0.5 %	0.15 [0.02, 1.33]
Kelly 1993 cb	0.2574 (0.9153)		0.7 %	1.29 [0.22, 7.78]
Kelly 1993 ss	-1.3571 (1.0345)		0.5 %	0.26 [0.03, 1.96]
Stall 1999@	0.2811 (0.5253)		2.1 %	1.32 [0.47, 3.71]
Roffman 1997	-0.4002 (0.2562)		8.9 %	0.67 [0.41, 1.11]
Choi 1996@	-0.2267 (0.6951)		1.2 %	0.80 [0.20, 3.11]
Rotheram-Borus 2001@	-0.2113 (0.8457)		0.8 %	0.81 [0.15, 4.25]
Kalichman 2001#	-1.1123 (0.5605)		1.9 %	0.33 [0.11, 0.99]
Harding 2004	-0.5509 (1.2099)		0.4 %	0.58 [0.05, 6.17]
Carballo-Diguez2005	-0.1466 (0.3349)		5.2 %	0.86 [0.45, 1.66]
Subtotal (95% CI)			**36.6 %**	**0.70 [0.55, 0.90]**

Heterogeneity: Chi² = 12.98, df = 17 (P = 0.74); I² =0.0%

Test for overall effect: Z = 2.82 (P = 0.0048)

0.2   0.5   1.0   2.0   5.0

Favours treatment   Favours control

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 1 Intervention vs minimal to no HIV prevention

Outcome: 1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]

Study or subgroup	log [Rate ratio] (SE)	Rate ratio IV,Fixed,95% CI	Weight	Rate ratio IV,Fixed,95% CI
2 Individual-level interventions (Rate ratios)				
Rosser 1990 v@	0.5596 (1.3428)		0.3 %	1.75 [0.13, 24.32]
Rosser 1990 c@	-0.4212 (1.3974)		0.3 %	0.66 [0.04, 10.15]
Picciano 2001#	-0.7084 (0.5547)		1.9 %	0.49 [0.17, 1.46]
Patterson 2003 b#	0.0252 (0.4937)		2.4 %	1.03 [0.39, 2.70]
Patterson 2003 c#	-0.2421 (0.4926)		2.4 %	0.78 [0.30, 2.06]
Patterson 2003 s#	0.06 (0.4774)		2.6 %	1.06 [0.42, 2.71]
Richardson 2004 g@	0.0079 (0.4651)		2.7 %	1.01 [0.41, 2.51]
Richardson 2004 L@	-0.3334 (0.4875)		2.5 %	0.72 [0.28, 1.86]
R-B 2004 in person@	0.2519 (1.1209)		0.5 %	1.29 [0.14, 11.57]
R-B 2004 phone@	0.4262 (1.2284)		0.4 %	1.53 [0.14, 17.01]
Healthy Living 2007#	-0.3354 (0.2201)		12.1 %	0.72 [0.46, 1.10]
Subtotal (95% CI)			**28.0 %**	**0.80 [0.60, 1.06]**

Heterogeneity: Chi² = 2.75, df = 10 (P = 0.99); I² =0.0%

Test for overall effect: Z = 1.54 (P = 0.12)

0.2   0.5   1.0   2.0   5.0

Favours treatment   Favours control

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 1 Intervention vs minimal to no HIV prevention

Outcome: 1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]

Study or subgroup	log [Rate ratio] (SE)	Rate ratio IV,Fixed,95% CI	Weight	Rate ratio IV,Fixed,95% CI
3 Community-level interventions (Rate ratios)				
Kelly 1991	-0.1137 (0.3816)		4.0 %	0.89 [0.42, 1.89]
Kelly 1997	-1.3736 (0.343)		5.0 %	0.25 [0.13, 0.50]
Kegeles 1996@	-0.343 (1.0907)		0.5 %	0.71 [0.08, 6.02]
CDC ACDP 1999@	-0.1919 (0.3692)		4.3 %	0.83 [0.40, 1.70]
Hoff 1997	-0.1025 (0.317)		5.8 %	0.90 [0.48, 1.68]
Shepherd 1997	-0.0572 (0.8487)		0.8 %	0.94 [0.18, 4.98]
Miller 1998	-0.4789 (0.4172)		3.4 %	0.62 [0.27, 1.40]
Elford 2001@	0.1169 (0.3676)		4.3 %	1.12 [0.55, 2.31]
Flowers 2002	-0.3185 (0.2998)		6.5 %	0.73 [0.40, 1.31]
Kegeles 2002@	-0.413 (0.9601)		0.6 %	0.66 [0.10, 4.34]
Amirkhanian 2005@	-0.4654 (1.6351)		0.2 %	0.63 [0.03, 15.48]
Subtotal (95% CI)			**35.4 %**	**0.70 [0.55, 0.91]**

Heterogeneity: Chi^2 = 11.93, df = 10 (P = 0.29); I^2 =16%

Test for overall effect: Z = 2.74 (P = 0.0062)

0.2 0.5 1.0 2.0 5.0

Favours treatment   Favours control

Analysis 1.2. Comparison 1 Intervention vs minimal to no HIV prevention, Outcome 2 PR Proportion reporting any unprotected sex [# = RR substituted for PR].

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 1 Intervention vs minimal to no HIV prevention

Outcome: 2 PR Proportion reporting any unprotected sex [# = RR substituted for PR]

Study or subgroup	log [Prevalence ratio] (SE)	Prevalence ratio IV,Fixed,95% CI	Weight	Prevalence ratio IV,Fixed,95% CI
1 Small group interventions (Prevalence ratios)				
Kelly 1989#	-0.9396 (0.27)		2.0 %	0.39 [0.23, 0.66]
Coates 1989@	0 (0.6614)		0.3 %	1.00 [0.27, 3.66]
Rosser 1990 e@	0.2719 (0.9951)		0.1 %	1.31 [0.19, 9.23]
Rosser 1990 s@	0.7828 (1.0131)		0.1 %	2.19 [0.30, 15.93]
Tudiver 1992 p1@	-0.3642 (0.2674)		2.0 %	0.69 [0.41, 1.17]
Tudiver 1992 c4@	-0.0075 (0.2955)		1.7 %	0.99 [0.56, 1.77]
Roffman 1998#	-0.3626 (0.1501)		6.5 %	0.70 [0.52, 0.93]
Peterson 1996 3s	-0.6506 (0.4327)		0.8 %	0.52 [0.22, 1.22]
Peterson 1996 1s	0 (0.4356)		0.8 %	1.00 [0.43, 2.35]
Kelly 1993 cb	-0.0451 (0.886)		0.2 %	0.96 [0.17, 5.43]
Kelly 1993 ss	-0.8409 (1.171)		0.1 %	0.43 [0.04, 4.28]
Stall 1999@	0.2811 (0.3464)		1.2 %	1.32 [0.67, 2.61]
Roffman 1997	-0.2377 (0.1443)		7.0 %	0.79 [0.59, 1.05]
Choi 1996@	-0.2267 (0.2024)		3.6 %	0.80 [0.54, 1.19]
Rotheram-Borus 2001@	-0.2113 (0.3816)		1.0 %	0.81 [0.38, 1.71]
Kalichman 2001#	-1.1123 (0.286)		1.8 %	0.33 [0.19, 0.58]
Harding 2004	-0.7449 (1.054)		0.1 %	0.47 [0.06, 3.75]
Carballo-Diguez2005	-0.0272 (0.2256)		2.9 %	0.97 [0.63, 1.51]
Subtotal (95% CI)			**32.3 %**	**0.73 [0.64, 0.84]**
Heterogeneity: Chi² = 22.81, df = 17 (P = 0.16); I² =25%				
Test for overall effect: Z = 4.58 (P < 0.00001)				
2 Individual-level interventions (Prevalence ratios)				
Rosser 1990 v@	0.5596 (0.9965)		0.1 %	1.75 [0.25, 12.34]
Rosser 1990 c@	-0.4212 (1.0769)		0.1 %	0.66 [0.08, 5.42]
Picciano 2001#	-0.7084 (0.2871)		1.8 %	0.49 [0.28, 0.86]

0.2   0.5   1.0   2.0   5.0

Favours treatment   Favours control

(*Continued* . . .)

Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men (Review)   **62**

Copyright © 2009 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Study or subgroup	log [Prevalence ratio] (SE)	Prevalence ratio IV,Fixed,95% CI	Weight	Prevalence ratio IV,Fixed,95% CI
Patterson 2003 b#	0.0252 (0.2077)		3.4 %	1.03 [0.68, 1.54]
Patterson 2003 c#	-0.2421 (0.2097)		3.3 %	0.78 [0.52, 1.18]
Patterson 2003 s#	0.06 (0.2082)		3.4 %	1.06 [0.71, 1.60]
Richardson 2004 g@	0.0079 (0.2529)		2.3 %	1.01 [0.61, 1.65]
Richardson 2004 L@	-0.3334 (0.2636)		2.1 %	0.72 [0.43, 1.20]
R-B 2004 in person@	0.2519 (0.3296)		1.3 %	1.29 [0.67, 2.45]
R-B 2004 phone@	0.4262 (0.3164)		1.5 %	1.53 [0.82, 2.85]
Healthy Living 2007#	-0.3354 (0.1147)		11.1 %	0.72 [0.57, 0.90]
Subtotal (95% CI)			**30.4 %**	**0.84 [0.74, 0.97]**

Heterogeneity: Chi² = 14.48, df = 10 (P = 0.15); I² =31%

Test for overall effect: Z = 2.46 (P = 0.014)

3 Community-level interventions (Prevalence ratios)

Kelly 1991	-0.3166 (0.2138)		3.2 %	0.73 [0.48, 1.11]
Kelly 1997	-0.5366 (0.2151)		3.2 %	0.58 [0.38, 0.89]
Kegeles 1996@	-0.343 (0.2443)		2.4 %	0.71 [0.44, 1.15]
CDC ACDP 1999@	-0.1919 (0.3001)		1.6 %	0.83 [0.46, 1.49]
Hoff 1997	-0.1025 (0.1833)		4.3 %	0.90 [0.63, 1.29]
Shepherd 1997	-0.3365 (0.7999)		0.2 %	0.71 [0.15, 3.43]
Miller 1998	-0.2675 (0.2489)		2.4 %	0.77 [0.47, 1.25]
Elford 2001@	0.1169 (0.2292)		2.8 %	1.12 [0.72, 1.76]
Flowers 2002	-0.155 (0.163)		5.5 %	0.86 [0.62, 1.18]
Kegeles 2002@	-0.413 (0.2842)		1.8 %	0.66 [0.38, 1.15]
Amirkhanian 2005@	-0.4654 (0.1217)		9.9 %	0.63 [0.49, 0.80]
Subtotal (95% CI)			**37.3 %**	**0.75 [0.66, 0.84]**

Heterogeneity: Chi² = 8.64, df = 10 (P = 0.57); I² =0.0%

Test for overall effect: Z = 4.66 (P < 0.00001)

| **Total (95% CI)** | | | **100.0 %** | **0.77 [0.72, 0.83]** |

Heterogeneity: Chi² = 48.36, df = 39 (P = 0.14); I² =19%

Test for overall effect: Z = 6.81 (P < 0.00001)

Test for subgroup differences: Chi² = 2.44, df = 2 (P = 0.30), I² =18%

```
         0.2    0.5   1.0   2.0   5.0
       Favours treatment    Favours control
```

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 1 Intervention vs minimal to no HIV prevention

Outcome: 2 PR Proportion reporting any unprotected sex [# = RR substituted for PR]

Study or subgroup	log [Prevalence ratio] (SE)	Prevalence ratio IV,Fixed,95% CI	Weight	Prevalence ratio IV,Fixed,95% CI
1 Small group interventions (Prevalence ratios)				
Kelly 1989#	-0.9396 (0.27)		2.0 %	0.39 [0.23, 0.66]
Coates 1989@	0 (0.6614)		0.3 %	1.00 [0.27, 3.66]
Rosser 1990 e@	0.2719 (0.9951)		0.1 %	1.31 [0.19, 9.23]
Rosser 1990 s@	0.7828 (1.0131)		0.1 %	2.19 [0.30, 15.93]
Tudiver 1992 p1@	-0.3642 (0.2674)		2.0 %	0.69 [0.41, 1.17]
Tudiver 1992 c4@	-0.0075 (0.2955)		1.7 %	0.99 [0.56, 1.77]
Roffman 1998#	-0.3626 (0.1501)		6.5 %	0.70 [0.52, 0.93]
Peterson 1996 3s	-0.6506 (0.4327)		0.8 %	0.52 [0.22, 1.22]
Peterson 1996 1s	0 (0.4356)		0.8 %	1.00 [0.43, 2.35]
Kelly 1993 cb	-0.0451 (0.886)		0.2 %	0.96 [0.17, 5.43]
Kelly 1993 ss	-0.8409 (1.171)		0.1 %	0.43 [0.04, 4.28]
Stall 1999@	0.2811 (0.3464)		1.2 %	1.32 [0.67, 2.61]
Roffman 1997	-0.2377 (0.1443)		7.0 %	0.79 [0.59, 1.05]
Choi 1996@	-0.2267 (0.2024)		3.6 %	0.80 [0.54, 1.19]
Rotheram-Borus 2001@	-0.2113 (0.3816)		1.0 %	0.81 [0.38, 1.71]
Kalichman 2001#	-1.1123 (0.286)		1.8 %	0.33 [0.19, 0.58]
Harding 2004	-0.7449 (1.054)		0.1 %	0.47 [0.06, 3.75]
Carballo-Diguez2005	-0.0272 (0.2256)		2.9 %	0.97 [0.63, 1.51]
Subtotal (95% CI)			**32.3 %**	**0.73 [0.64, 0.84]**

Heterogeneity: Chi² = 22.81, df = 17 (P = 0.16); I² =25%

Test for overall effect: Z = 4.58 (P < 0.00001)

0.2   0.5   1.0   2.0   5.0

Favours treatment   Favours control

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 1 Intervention vs minimal to no HIV prevention

Outcome: 2 PR Proportion reporting any unprotected sex [# = RR substituted for PR]

Study or subgroup	log [Prevalence ratio] (SE)	Prevalence ratio IV,Fixed,95% CI	Weight	Prevalence ratio IV,Fixed,95% CI
2 Individual-level interventions (Prevalence ratios)				
Rosser 1990 v@	0.5596 (0.9965)		0.1 %	1.75 [0.25, 12.34]
Rosser 1990 c@	-0.4212 (1.0769)		0.1 %	0.66 [0.08, 5.42]
Picciano 2001#	-0.7084 (0.2871)		1.8 %	0.49 [0.28, 0.86]
Patterson 2003 b#	0.0252 (0.2077)		3.4 %	1.03 [0.68, 1.54]
Patterson 2003 c#	-0.2421 (0.2097)		3.3 %	0.78 [0.52, 1.18]
Patterson 2003 s#	0.06 (0.2082)		3.4 %	1.06 [0.71, 1.60]
Richardson 2004 g@	0.0079 (0.2529)		2.3 %	1.01 [0.61, 1.65]
Richardson 2004 L@	-0.3334 (0.2636)		2.1 %	0.72 [0.43, 1.20]
R-B 2004 in person@	0.2519 (0.3296)		1.3 %	1.29 [0.67, 2.45]
R-B 2004 phone@	0.4262 (0.3164)		1.5 %	1.53 [0.82, 2.85]
Healthy Living 2007#	-0.3354 (0.1147)		11.1 %	0.72 [0.57, 0.90]
Subtotal (95% CI)			**30.4 %**	**0.84 [0.74, 0.97]**
Heterogeneity: Chi² = 14.48, df = 10 (P = 0.15); I² =31%				
Test for overall effect: Z = 2.46 (P = 0.014)				

0.2   0.5   1.0   2.0   5.0

Favours treatment   Favours control

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 1 Intervention vs minimal to no HIV prevention

Outcome: 2 PR Proportion reporting any unprotected sex [# = RR substituted for PR]

Study or subgroup	log [Prevalence ratio] (SE)	Prevalence ratio IV,Fixed,95% CI	Weight	Prevalence ratio IV,Fixed,95% CI
3 Community-level interventions (Prevalence ratios)				
Kelly 1991	-0.3166 (0.2138)		3.2 %	0.73 [0.48, 1.11]
Kelly 1997	-0.5366 (0.2151)		3.2 %	0.58 [0.38, 0.89]
Kegeles 1996@	-0.343 (0.2443)		2.4 %	0.71 [0.44, 1.15]
CDC ACDP 1999@	-0.1919 (0.3001)		1.6 %	0.83 [0.46, 1.49]
Hoff 1997	-0.1025 (0.1833)		4.3 %	0.90 [0.63, 1.29]
Shepherd 1997	-0.3365 (0.7999)		0.2 %	0.71 [0.15, 3.43]
Miller 1998	-0.2675 (0.2489)		2.4 %	0.77 [0.47, 1.25]
Elford 2001@	0.1169 (0.2292)		2.8 %	1.12 [0.72, 1.76]
Flowers 2002	-0.155 (0.163)		5.5 %	0.86 [0.62, 1.18]
Kegeles 2002@	-0.413 (0.2842)		1.8 %	0.66 [0.38, 1.15]
Amirkhanian 2005@	-0.4654 (0.1217)		9.9 %	0.63 [0.49, 0.80]
Subtotal (95% CI)			**37.3 %**	**0.75 [0.66, 0.84]**
Heterogeneity: Chi² = 8.64, df = 10 (P = 0.57); I² =0.0%				
Test for overall effect: Z = 4.66 (P < 0.00001)				

0.2   0.5   1.0   2.0   5.0

Favours treatment   Favours control

Analysis 2.1. Comparison 2 Experimental vs standard or other HIV prevention, Outcome 1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR].

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 2 Experimental vs standard or other HIV prevention

Outcome: 1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]

Study or subgroup	log [Rate ratio] (SE)	Rate ratio IV,Fixed,95% CI	Weight	Rate ratio IV,Fixed,95% CI
1 Small group interventions (Rate Ratios)				
Valdiserri 1989*	-0.3442 (0.2964)		5.2 %	0.71 [0.40, 1.27]
Cleary 1995*@	-0.1033 (0.6034)		1.3 %	0.90 [0.28, 2.94]
Imrie 2001*@	-0.1438 (0.3404)		4.0 %	0.87 [0.44, 1.69]
Rosser 2002*@	-0.47 (0.5104)		1.8 %	0.63 [0.23, 1.70]
Sampaio 2002*@	-0.4568 (1.0179)		0.4 %	0.63 [0.09, 4.66]
Shoptaw 2005 cbt*	0.0592 (0.3899)		3.0 %	1.06 [0.49, 2.28]
Shoptaw 2005 g*	-0.8142 (0.3872)		3.1 %	0.44 [0.21, 0.95]
Wolitski 2005*	-0.0989 (0.2693)		6.3 %	0.91 [0.53, 1.54]
Subtotal (95% CI)			**25.0 %**	**0.77 [0.59, 1.01]**
Heterogeneity: Chi² = 3.55, df = 7 (P = 0.83); I² =0.0%				
Test for overall effect: Z = 1.91 (P = 0.056)				
2 Individual-level interventions (Rate Ratios)				
Gold 1995 sj*	-0.2163 (0.8649)		0.6 %	0.81 [0.15, 4.39]
Gold 1995 po*	0.0322 (0.7818)		0.7 %	1.03 [0.22, 4.78]
Gold 1998 sjp*	0.1057 (0.9193)		0.5 %	1.11 [0.18, 6.74]
Gold 1998 rse*	0.0317 (0.9223)		0.5 %	1.03 [0.17, 6.29]
Sorensen 2003*	-0.4227 (0.7403)		0.8 %	0.66 [0.15, 2.80]
Dilley 2002 n*@	-0.7725 (0.5316)		1.6 %	0.46 [0.16, 1.31]
Dilley 2002 d*@	-0.47 (0.5381)		1.6 %	0.63 [0.22, 1.79]
Explore 2004*	-0.1008 (0.0875)		59.9 %	0.90 [0.76, 1.07]
Read 2006*#	-0.2043 (0.5927)		1.3 %	0.82 [0.26, 2.60]
Dilley 2007*#	-0.4055 (0.2503)		7.3 %	0.67 [0.41, 1.09]
Subtotal (95% CI)			**75.0 %**	**0.86 [0.73, 1.00]**
Heterogeneity: Chi² = 3.40, df = 9 (P = 0.95); I² =0.0%				
Test for overall effect: Z = 1.99 (P = 0.047)				
Total (95% CI)			**100.0 %**	**0.83 [0.73, 0.95]**
Heterogeneity: Chi² = 7.38, df = 17 (P = 0.98); I² =0.0%				
Test for overall effect: Z = 2.68 (P = 0.0075)				
Test for subgroup differences: Chi² = 0.43, df = 1 (P = 0.51), I² =0.0%				

0.2   0.5   1.0   2.0   5.0

Favours treatment   Favours control

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 2 Experimental vs standard or other HIV prevention

Outcome: 1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]

Study or subgroup	log [Rate ratio] (SE)	Rate ratio IV,Fixed,95% CI	Weight	Rate ratio IV,Fixed,95% CI
1 Small group interventions (Rate Ratios)				
Valdiserri 1989*	-0.3442 (0.2964)		5.2 %	0.71 [0.40, 1.27]
Cleary 1995*@	-0.1033 (0.6034)		1.3 %	0.90 [0.28, 2.94]
Imrie 2001*@	-0.1438 (0.3404)		4.0 %	0.87 [0.44, 1.69]
Rosser 2002*@	-0.47 (0.5104)		1.8 %	0.63 [0.23, 1.70]
Sampaio 2002*@	-0.4568 (1.0179)		0.4 %	0.63 [0.09, 4.66]
Shoptaw 2005 cbt*	0.0592 (0.3899)		3.0 %	1.06 [0.49, 2.28]
Shoptaw 2005 g*	-0.8142 (0.3872)		3.1 %	0.44 [0.21, 0.95]
Wolitski 2005*	-0.0989 (0.2693)		6.3 %	0.91 [0.53, 1.54]
Subtotal (95% CI)			**25.0 %**	**0.77 [0.59, 1.01]**

Heterogeneity: Chi² = 3.55, df = 7 (P = 0.83); I² =0.0%

Test for overall effect: Z = 1.91 (P = 0.056)

0.2 0.5 1.0 2.0 5.0

Favours treatment   Favours control

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 2 Experimental vs standard or other HIV prevention

Outcome: 1 RR Occasions of or partners for unprotected sex [@ = PR substituted for RR]

Study or subgroup	log [Rate ratio] (SE)	Rate ratio IV,Fixed,95% CI	Weight	Rate ratio IV,Fixed,95% CI
2 Individual-level interventions (Rate Ratios)				
Gold 1995 sj*	-0.2163 (0.8649)		0.6 %	0.81 [0.15, 4.39]
Gold 1995 po*	0.0322 (0.7818)		0.7 %	1.03 [0.22, 4.78]
Gold 1998 sjp*	0.1057 (0.9193)		0.5 %	1.11 [0.18, 6.74]
Gold 1998 rse*	0.0317 (0.9223)		0.5 %	1.03 [0.17, 6.29]
Sorensen 2003*	-0.4227 (0.7403)		0.8 %	0.66 [0.15, 2.80]
Dilley 2002 n*@	-0.7725 (0.5316)		1.6 %	0.46 [0.16, 1.31]
Dilley 2002 d*@	-0.47 (0.5381)		1.6 %	0.63 [0.22, 1.79]
Explore 2004*	-0.1008 (0.0875)		59.9 %	0.90 [0.76, 1.07]
Read 2006*#	-0.2043 (0.5927)		1.3 %	0.82 [0.26, 2.60]
Dilley 2007*#	-0.4055 (0.2503)		7.3 %	0.67 [0.41, 1.09]
Subtotal (95% CI)			**75.0 %**	**0.86 [0.73, 1.00]**

Heterogeneity: Chi² = 3.40, df = 9 (P = 0.95); I² =0.0%

Test for overall effect: Z = 1.99 (P = 0.047)

```
0.2   0.5   1.0   2.0   5.0
```
Favours treatment   Favours control

Analysis 2.2. Comparison 2 Experimental vs standard or other HIV prevention, Outcome 2 PR Proportion reporting any unprotected sex [# = RR substituted for PR].

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 2 Experimental vs standard or other HIV prevention

Outcome: 2 PR Proportion reporting any unprotected sex [# = RR substituted for PR]

Study or subgroup	log [Prevalence ratio] (SE)	Prevalence ratio IV,Fixed,95% CI	Weight	Prevalence ratio IV,Fixed,95% CI
1 Small group interventions (Prevalence ratios)				
Valdiserri 1989*	-0.1484 (0.1162)		4.0 %	0.86 [0.69, 1.08]
Cleary 1995*@	-0.1033 (0.4412)		0.3 %	0.90 [0.38, 2.14]
Imrie 2001*@	-0.1438 (0.1182)		3.8 %	0.87 [0.69, 1.09]
Rosser 2002*@	-0.47 (0.3832)		0.4 %	0.63 [0.29, 1.32]
Sampaio 2002*@	-0.4568 (1.2175)		0.0 %	0.63 [0.06, 6.89]
Shoptaw 2005 cbt*	-0.087 (0.2263)		1.0 %	0.92 [0.59, 1.43]
Shoptaw 2005 g*	0.1203 (0.1855)		1.6 %	1.13 [0.78, 1.62]
Wolitski 2005*	-0.1933 (0.1034)		5.0 %	0.82 [0.67, 1.01]
Subtotal (95% CI)			**16.0 %**	**0.87 [0.78, 0.97]**
Heterogeneity: Chi² = 3.11, df = 7 (P = 0.87); I² =0.0%				
Test for overall effect: Z = 2.40 (P = 0.016)				
2 Individual-level interventions (Prevalence ratios)				
Gold 1995 sj*	-0.1001 (0.254)		0.8 %	0.90 [0.55, 1.49]
Gold 1995 po*	0.0636 (0.2369)		1.0 %	1.07 [0.67, 1.70]
Gold 1998 sjp*	0.1525 (0.2584)		0.8 %	1.16 [0.70, 1.93]
Gold 1998 rse*	0.0317 (0.2682)		0.7 %	1.03 [0.61, 1.75]
Sorensen 2003*	-0.1516 (0.1763)		1.7 %	0.86 [0.61, 1.21]
Dilley 2002 n*@	-0.7725 (0.3224)		0.5 %	0.46 [0.25, 0.87]
Dilley 2002 d*@	-0.47 (0.2951)		0.6 %	0.63 [0.35, 1.11]
Explore 2004*	-0.0409 (0.0267)		74.9 %	0.96 [0.91, 1.01]
Read 2006*#	-0.0816 (0.3695)		0.4 %	0.92 [0.45, 1.90]
Dilley 2007*#	-0.4055 (0.1466)		2.5 %	0.67 [0.50, 0.89]
Subtotal (95% CI)			**84.0 %**	**0.94 [0.90, 0.99]**
Heterogeneity: Chi² = 14.24, df = 9 (P = 0.11); I² =37%				
Test for overall effect: Z = 2.33 (P = 0.020)				
Total (95% CI)			**100.0 %**	**0.93 [0.89, 0.97]**
Heterogeneity: Chi² = 18.96, df = 17 (P = 0.33); I² =10%				
Test for overall effect: Z = 3.09 (P = 0.0020)				
Test for subgroup differences: Chi² = 1.61, df = 1 (P = 0.20), I² =38%				

0.5   0.7   1.0   1.5   2.0

Favours treatment   Favours control

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 2 Experimental vs standard or other HIV prevention

Outcome: 2 PR Proportion reporting any unprotected sex [# = RR substituted for PR]

Study or subgroup	log [Prevalence ratio] (SE)	Prevalence ratio IV,Fixed,95% CI	Weight	Prevalence ratio IV,Fixed,95% CI
1 Small group interventions (Prevalence ratios)				
Valdiserri 1989*	-0.1484 (0.1162)		4.0 %	0.86 [0.69, 1.08]
Cleary 1995*@	-0.1033 (0.4412)		0.3 %	0.90 [0.38, 2.14]
Imrie 2001*@	-0.1438 (0.1182)		3.8 %	0.87 [0.69, 1.09]
Rosser 2002*@	-0.47 (0.3832)		0.4 %	0.63 [0.29, 1.32]
Sampaio 2002*@	-0.4568 (1.2175)		0.0 %	0.63 [0.06, 6.89]
Shoptaw 2005 cbt*	-0.087 (0.2263)		1.0 %	0.92 [0.59, 1.43]
Shoptaw 2005 g*	0.1203 (0.1855)		1.6 %	1.13 [0.78, 1.62]
Wolitski 2005*	-0.1933 (0.1034)		5.0 %	0.82 [0.67, 1.01]
Subtotal (95% CI)			**16.0 %**	**0.87 [0.78, 0.97]**

Heterogeneity: Chi² = 3.11, df = 7 (P = 0.87); I² =0.0%

Test for overall effect: Z = 2.40 (P = 0.016)

0.5   0.7   1.0   1.5   2.0

Favours treatment   Favours control

Review: Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men

Comparison: 2 Experimental vs standard or other HIV prevention

Outcome: 2 PR Proportion reporting any unprotected sex [# = RR substituted for PR]

Study or subgroup	log [Prevalence ratio] (SE)	Prevalence ratio IV,Fixed,95% CI	Weight	Prevalence ratio IV,Fixed,95% CI
2 Individual-level interventions (Prevalence ratios)				
Gold 1995 sj*	-0.1001 (0.254)		0.8 %	0.90 [0.55, 1.49]
Gold 1995 po*	0.0636 (0.2369)		1.0 %	1.07 [0.67, 1.70]
Gold 1998 sjp*	0.1525 (0.2584)		0.8 %	1.16 [0.70, 1.93]
Gold 1998 rse*	0.0317 (0.2682)		0.7 %	1.03 [0.61, 1.75]
Sorensen 2003*	-0.1516 (0.1763)		1.7 %	0.86 [0.61, 1.21]
Dilley 2002 n*@	-0.7725 (0.3224)		0.5 %	0.46 [0.25, 0.87]
Dilley 2002 d*@	-0.47 (0.2951)		0.6 %	0.63 [0.35, 1.11]
Explore 2004*	-0.0409 (0.0267)		74.9 %	0.96 [0.91, 1.01]
Read 2006*#	-0.0816 (0.3695)		0.4 %	0.92 [0.45, 1.90]
Dilley 2007*#	-0.4055 (0.1466)		2.5 %	0.67 [0.50, 0.89]
Subtotal (95% CI)			**84.0 %**	**0.94 [0.90, 0.99]**
Heterogeneity: Chi² = 14.24, df = 9 (P = 0.11); I² =37%				
Test for overall effect: Z = 2.33 (P = 0.020)				

```
        0.5   0.7   1.0   1.5   2.0
        Favours treatment    Favours control
```

WHAT'S NEW

Last assessed as up-to-date: 30 April 2008.

14 May 2008	Amended	Converted to new review format.
7 February 2008	New citation required and conclusions have changed	Substantive amendment

HISTORY

Protocol first published: Issue 3, 1998

Review first published: Issue 1, 2003

CONTRIBUTIONS OF AUTHORS

WDJ - eligibility criteria, search strategies, location of studies, review for eligibility, data abstraction, calculation of effects, analysis

All reviewers contributed to conceptualization, interpretation, and writing.

DECLARATIONS OF INTEREST

Three authors work in HIV prevention and one (WDJ) worked on one of the included studies [CDC ACDP 1999@].

SOURCES OF SUPPORT

Internal sources

- Centers for Disease Control and Prevention (CDC), USA.
- Emory University Rollins School of Public Health, USA.
- Florida International University, USA.
- San Francisco State University, USA.
- Johns Hopkins Bloomberg School of Public Health, USA.

External sources

- No sources of support supplied

INDEX TERMS

Medical Subject Headings (MeSH)

Health Knowledge, Attitudes, Practice; HIV Infections [*prevention & control; transmission]; *Homosexuality, Male; Randomized Controlled Trials as Topic; Risk-Taking; Safe Sex; Sexually Transmitted Diseases [prevention & control]

MeSH check words

Humans; Male