



Practice of Epidemiology

Estimating the Efficacy of Preexposure Prophylaxis for HIV Prevention Among Participants With a Threshold Level of Drug Concentration

James Y. Dai*, Peter B. Gilbert, James P. Hughes, and Elizabeth R. Brown

* Correspondence to Dr. James Y. Dai, M2-C200, Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, WA 98109 (e-mail: jdai@fhcrc.org).

Initially submitted January 26, 2012; accepted for publication May 10, 2012.

Assays for detecting levels of antiretroviral drugs in study participants are increasingly popular in preexposure prophylaxis (PrEP) trials, since they provide an objective measure of adherence. Current correlation analyses of drug concentration data are prone to bias. In this article, we formulate the causal estimand of prevention efficacy among drug compliers, those who would have had a threshold level of drug concentration had they been assigned to the drug arm of the trial. The identifiability of the causal estimand is facilitated by exploiting the exclusion restriction; that is, drug noncompliers do not acquire any prevention benefit. In addition, we develop an approach to sensitivity analysis that relaxes the exclusion restriction. Applications to published data from 2 PrEP trials, namely the Preexposure Prophylaxis Initiative (iPrEx) trial and the Centre for the AIDS Programme of Research in South Africa (CAPRISA) 004 trial, suggest high efficacy estimates among drug compliers (in the iPrEx trial, odds ratio = 0.097 (95% confidence interval: 0.027, 0.352); in the CAPRISA 004 trial, odds ratio = 0.104 (95% confidence interval: 0.024, 0.447)). In summary, the proposed inferential method provides an unbiased assessment of PrEP efficacy among drug compliers, thus adding to the primary intention-to-treat analysis and correlation analyses of drug concentration data.

causal inference; compliance; exclusion restriction; potential outcome; principal stratification; two-phase sampling

Abbreviations: CACE, complier-average causal effect; CAPRISA, Centre for the AIDS Programme of Research in South Africa; CI, confidence interval; HIV, human immunodeficiency virus; iPrEx, Preexposure Prophylaxis Initiative; ITT, intention-to-treat; PrEP, preexposure prophylaxis.

Preexposure prophylaxis (PrEP) for human immunodeficiency virus (HIV) prevention is a promising strategy that is currently being tested globally using antiretroviral agents delivered either orally or topically prior to sexual exposure. In the Preexposure Prophylaxis Initiative (iPrEx) trial, where men who have sex with men were randomized to receive either daily emtricitabine–tenofovir disoproxil fumarate pills or placebo pills, the intention-to-treat (ITT) analysis found a 44% reduction in HIV incidence in the active drug arm as compared with the placebo arm (1). In the Centre for the AIDS Programme of Research in South Africa (CAPRISA) 004 trial, coitus-related use of a tenofovir 1% vaginal gel was shown to reduce HIV acquisition by 39% among high-risk South African women (2). Several other clinical trials (either recently completed or

still ongoing) are evaluating or have evaluated tenofovir-based PrEP regimens in different risk groups, such as heterosexuals and injection drug users (3).

Adherence to the daily or coitus-related regimen is critical to PrEP effectiveness. Both the iPrEx trial and the CAPRISA 004 trial found that study participants who used PrEP consistently were less susceptible to HIV infection. Precise measurement of drug adherence, however, has been a major challenge in PrEP studies. Self-reported adherence has often been overreported (1, 4). Several more objective measures of adherence have been used in PrEP studies, including pill counts, counts of used gel applicators, and measurements of drug concentrations in participants' blood, tissue, or hair. In the iPrEx trial, for instance, a nested case-control study was conducted to detect the presence of the

drug in plasma and in peripheral-blood mononuclear cells at the infection visit for cases and for HIV-negative controls matched by study visit and site (1). No drug was detected among participants in the placebo arm, as expected. Remarkably, comparison of the HIV risks between participants with a detectable level of drug concentration and participants without a detectable level yielded a 92% reduction in HIV risk (1). In the CAPRISA 004 trial, tenofovir concentrations in undiluted aspirated cervicovaginal fluid were assessed for women in the tenofovir gel arm (5). Drug assays were performed using samples collected at the first visit postinfection for seroconverters and using samples from a randomly selected visit for uninfected women. Results suggested that women with a tenofovir concentration greater than 1,000 ng/mL had a 74% reduction in HIV incidence compared with women in the placebo group (2.4 per 100 person-years vs. 9.1 per 100 person-years; incidence rate ratio = 0.26; $P = 0.01$).

All of these analyses attempted to assess PrEP efficacy among compliers using drug concentration as a marker of adherence. Formally, we define “prevention efficacy” in PrEP trials as the treatment effect from an antiretroviral drug among persons who comply with the assigned regimen. It describes the biological effect of the drug, while “prevention effectiveness,” the average treatment effect estimated by means of ITT analysis, depends on both the biological effect of the drug and compliance with the regimen. Drug concentration assays provide a biomarker proxy for adherence, which is more accurate than self-reported adherence, but they are available only for persons in the drug arm of a trial. Correlation of drug concentrations and HIV infections in the drug arm (e.g., the iPrEx analysis) is prone to selection bias, since compliers and noncompliers may have different health conditions and different HIV risk profiles. Comparison of compliers in the drug arm with the placebo recipients is not protected by randomization either. Consequently, unless strong and unverifiable assumptions are made, none of the aforementioned drug-level analyses yield an effect estimate that is interpretable as being due to PrEP use, since the comparisons are not conducted between the randomized drug and placebo assignment groups.

Using the potential outcomes framework in causal inference (6–9), we present here an estimation procedure for assessing the prevention efficacy of a PrEP regimen among drug compliers, defined as persons who would have achieved a threshold level of drug concentration had they been assigned to the drug arm. To unify notation in the two trials, we define a threshold level of drug concentration as the level above which investigators hypothesize that the drug would protect against HIV infection if the drug were efficacious. In the iPrEx analysis of the daily pill-taking regimen, the threshold is any detectable level of the drug in the sampled blood. In the CAPRISA 004 analysis, this threshold level is defined as a cervicovaginal fluid tenofovir concentration greater than 1,000 ng/mL. We note that having a drug concentration above the threshold level is driven mostly by adherence, but it is also affected by the dosing regimen and individual genetic variation in pharmacological properties and host-cell biology (5). For the daily regimen prescribed in the iPrEx trial, for example, drug levels are probably

detectable if a tablet was taken in the last 14 days. For the coitally dependent use of gel in the CAPRISA 004 trial, the value of drug levels as a marker of adherence depends on the timing of the last sex act before a clinic visit.

Our causal estimand relates to the treatment effect among “would-be treatment compliers” studied in the compliance literature (10–12), defined as persons who adhere to the active treatment regimen, irrespective of their potential compliance status under the control treatment. This estimand is of interest in a superiority trial, double-blind or open-label, when the control is some basic and standard condition, or an inert placebo, so that compliance with the control regimen does not produce any effect. Another related estimand is the well-known complier-average causal effect (CACE) (13), the average effect of the treatment in the subpopulation of participants who would adhere to whatever regimen they were assigned. Often the CACE is assessed in open-label randomized clinical trials, where participants are classified into 4 types: compliers, defiers, never-takers, and always-takers, assuming that participants have access to the alternative treatment other than the one they are assigned to (13). In the double-blind PrEP trials we considered here, antiretroviral drugs seem to be well tolerated and there is no major safety concern, so blinding to treatment assignment is well kept. Therefore, similar to Rubin (12), we may assume that participants are comprised only of compliers and noncompliers; the latter group fails to comply under either assignment. To this end, our estimand can be viewed as a special form of CACE in double-blind trials.

The unique feature of the two trials we consider is that compliance in the drug arm is measured by an objective biomarker, rather than error-prone self-report questionnaires. The challenge in estimating prevention efficacy is that although drug compliers are directly identified in the drug arm, the drug compliers in the placebo arm are not identifiable. To achieve identifiability, we exploit the condition that drug noncompliers do not derive any protection against HIV infection, a version of the exclusion restriction widely known in the econometrics and statistics literature (14). Furthermore, we formulate a general sensitivity model that includes both situations where the exclusion restriction holds and situations where it does not hold. We develop a maximum likelihood estimation procedure accounting for the commonly used case-control sampling for drug assays and obtain prevention efficacy estimates for the iPrEx trial and the CAPRISA 004 trial. By varying the sensitivity parameter, we examine the impact of a nonzero protection effect in the drug noncompliers on estimating the prevention efficacy among drug compliers.

MATERIALS AND METHODS

Counterfactuals, principal stratification, and causal estimand

The counterfactual approach to causal analysis focuses on the collection of potential responses, say HIV infection status $Y(z)$, under all possible treatment options $Z = z$ (7, 8). In the context of PrEP trials, let us denote $Z = 1$ if a participant is assigned to the antiretroviral arm and $Z = 0$ if the

participant is assigned to the placebo arm. For a subject i , the counterfactuals contain a pair of potential outcomes $(Y_i(1), Y_i(0))$, among which only 1 potential outcome is realized, since in the trial participants receive only 1 treatment assignment. The causal effect of the treatment on HIV infection for this subject is based on the comparison between $Y_i(1)$ and $Y_i(0)$, say $Y_i(1) - Y_i(0)$, but the individual-level causal effect is not identifiable. Because the treatment assignment is randomized, the average $Y_i(z)$ in the trial can be estimated by the average Y among persons who actually receive the treatment assignment z , for $z = 0$ or $z = 1$, that is, $E[Y_i(z)] = E[Y_i(z)|z] = E[Y_i|z]$. Thus, the average causal treatment effect in the target population is identifiable.

A common use of the counterfactual framework is adjusting treatment effects for postrandomization variables, which are defined as those measured during follow-up, such as treatment noncompliance (13, 15, 16). Standard adjustment by means of regression modeling is susceptible to selection bias (17). One way to control for postrandomization variables is to compare values for the study endpoint in subgroups that share the same union of both sets of potential outcomes for the postrandomization variable, termed “principal strata” in the statistical literature (18). The key property of a principal stratum is that because it is based on potential outcomes every participant could have had, it is independent of treatment assignment and can be treated as a baseline covariate. Thus, the treatment effect estimated in a principal stratum is not affected by the postrandomization variable. We next illustrate the concept and the use of principal strata for the drug concentration data in the iPrEx trial and the CAPRISA 004 trial.

Define by $D(1)$ the potential outcome of whether the drug concentration would have been above the prespecified threshold level had a participant been assigned to the drug arm, and define by $Y(z)$ the potential outcome of the HIV infection had a participant been assigned to the z arm. Two principal strata are formed by $D(1)$: persons who have $D(1) = 1$ and those who have $D(1) = 0$. To ease the notation, we call the former stratum “drug compliers” and the latter stratum “drug non-compliers,” even though the definition of $D(1)$ is based solely on the potential drug concentration. As we discussed above, drug concentration serves as an approximate measure of adherence, as it is also affected by the dosing regimen and genetic variability of pharmacological parameters. This definition of compliers connects to the would-be treatment complier previously discussed in the compliance literature (10, 11). Compared with the well-known pill-taking complier strata in CACE (13), the principal strata here are only defined by $D(1)$ and thus are observable in the drug arm, but not $D(0)$ because the latter is identically zero for all placebo recipients. These properties are similar to principal strata in evaluating surrogates of protection in HIV vaccine trials (19, 20), where vaccine-induced immune responses are present only in the vaccine recipients.

A causal estimand measuring the efficacy of the PrEP regimen among drug compliers is defined as

$$\frac{\Pr(Y(1) = 1|D(1) = 1)}{\Pr(Y(0) = 1|D(1) = 1)},$$

which compares the HIV incidence among drug compliers in the active arm with the HIV incidence among those in the control arm, who would have been drug compliers had they been assigned to the active arm. This estimand is the ITT effect in the subgroup of drug compliers, addressing the question, For the subgroup that is assigned to the active treatment arm and complies with the dosing regimen, what is the average treatment efficacy when compared with the counterfactual scenario that they were assigned to the placebo arm? This is different from the question asked by the widely known CACE estimand, where a complier must adhere to whatever treatment (active or control) he or she is assigned. Indeed, if the active treatment does not create adverse effect and blinding is strictly maintained, compliance should not differ by arm; thus, the treatment effect among drug compliers is the CACE in this setting (12).

Neither of the aforementioned reported analyses of drug concentration data in the two trials yields a causal effect estimate. In the iPrEx trial, the analysis of the prophylaxis effect is essentially comparing HIV incidence among drug compliers in the drug arm with HIV incidence among drug noncompliers in the drug arm (1). In our notation, the estimand in that analysis is

$$\frac{\Pr(Y(1) = 1|D(1) = 1)}{\Pr(Y(1) = 1|D(1) = 0)}.$$

The conditioning on $z = 1$ is taken out both in the numerator and in the denominator because treatment assignment is independent of all potential outcomes. This estimand does not have a causal interpretation, because the contrast of HIV infection risk is not based on the same principal stratum. In other words, the participants who took their pills in the drug arm may systematically differ from the participants who did not take pills in the drug arm in terms of characteristics that are related to the risk of HIV infection. For example, drug compliers may have lower sexual risk-taking behavior. In the CAPRISA 004 trial, the comparison was conducted between the group of drug recipients who had tenofovir concentrations of more than 1,000 ng/mL and the placebo recipients, denoted in our notation by

$$\frac{\Pr(Y(1) = 1|D(1) = 1)}{\Pr(Y(0) = 1)},$$

which again is not a causal estimand for the same reason.

Identifiability, sensitivity parameter, and estimation

Drug compliers and drug noncompliers are not identifiable in the placebo arm. An experimental solution for future PrEP trials, if compliance across 2 arms is the same, would be to put an inert tracer into both the active product and the placebo, so that compliers in both arms could be identified. Analytically, we invoke a version of the exclusion restriction to facilitate identifiability (14). Contrary to typical self-reported adherence data, drug concentration as a proxy of adherence makes it defensible to invoke such an assumption, because it is a more accurate measure of

adherence. In the iPrEx trial, the absence of a detectable drug level is primarily caused by failure to have taken a pill in the past 14 days, and to some extent by the variability of individual pharmacokinetics. It seems plausible to assume that persons without a detectable level of the antiretroviral drug would be unlikely to receive any prevention benefit from the drug. To start, we make the following exclusion restriction assumption for the identifiability:

$$\Pr(Y(1) = 1|D(1) = 0) = \Pr(Y(0) = 1|D(1) = 0); \quad (1)$$

that is, those potential drug noncompliers do not derive any treatment effect. This is irrespective of their compliance status if assigned to the placebo arm. Because of the inert placebo being used, even if these persons switch from non-complying in the drug arm to complying in the placebo arm for some reason, they still do not obtain any benefit from the treatment.

Let the proportion of drug compliers be $\alpha = \Pr(D(1) = 1)$. Because of randomization, there will be the same proportions of drug compliers in the treatment arm and in the placebo arm. Thus, the observed HIV infection rate in the control arm, $\Pr(Y|Z = 0)$, is a mixture of 2 weighted probabilities:

$$\begin{aligned} \Pr(Y = 1|Z = 0) &= \Pr(Y(0) = 1) \\ &= \alpha\Pr(Y(0) = 1|D(1) = 1) + (1 - \alpha) \\ &\quad \times \Pr(Y(0) = 1|D(1) = 0). \end{aligned} \quad (2)$$

In the drug arm, both α and $\Pr(Y(1) = 1|D(1) = 0)$ are estimable, as is the infection rate in the placebo arm $\Pr(Y|Z = 0)$. From equations 1 and 2, we are now able to identify $\Pr(Y(0) = 1|D(1) = 1)$.

More generally, it is necessary to formulate a sensitivity model to assess the impact of deviating from the exclusion restriction. This is particularly useful for the CAPRISA 004 data, as 1,000 ng/mL was used as the threshold concentration. One can envision some residual protection remaining in those women with less than 1,000 ng/mL. Women could also show a low level of drug concentration because they have not engaged in recent sex acts. Even for the iPrEx trial, the exclusion restriction may not always be true for every participant, because the drug assay was conducted at only a single visit. A participant may have missed pills or gel use right before the visit but been an adherer all along, and thus may have received some degree of protection from the drug regimen. Similar approaches have been used for assessing the effect of preventive vaccines on postinfection outcomes in HIV vaccine trials (21, 22).

Consider the following logistic regression model for the association between treatment assignment, $D(1)$ as a baseline predictor, and infection status,

$$\begin{aligned} \log \left\{ \frac{\Pr(Y(z) = 1|z, D(1))}{\Pr(Y(z) = 0|z, D(1))} \right\} &= \beta_0 + \delta z + \beta_1 D(1) \\ &\quad + \beta_2 z D(1). \end{aligned}$$

Additional covariates in sexual risk-taking can be included to adjust for differential exposures between drug compliers and drug noncompliers. In this model, δ is the sensitivity parameter which captures the treatment effect among drug noncompliers, as we elaborate below. The parameter β_1 quantifies the difference in HIV risks between compliers and noncompliers in the placebo arm, a direct measure of selection bias when making inference on the 2 groups. The parameter of primary interest is β_2 , the log odds ratio between treatment and control among persons who have $D(1) = 1$,

$$\begin{aligned} &\frac{\Pr(Y = 1|Z = 1, D(1) = 1)/\Pr(Y = 0|Z = 1, D(1) = 1)}{\Pr(Y = 1|Z = 0, D(1) = 1)/\Pr(Y = 0|Z = 0, D(1) = 1)} \\ &= \frac{\Pr(Y(1) = 1|D(1) = 1)/\Pr(Y(1) = 0|D(1) = 1)}{\Pr(Y(0) = 1|D(1) = 1)/\Pr(Y(0) = 0|D(1) = 1)}. \end{aligned}$$

This approximates the relative risk, since HIV infection is generally a rare event. The set of parameters ($\beta_0, \delta, \beta_1, \beta_2$) is not identified, since we do not observe $D(1)$ for everyone. However, if we fix δ to a value in a user-specified, scientifically plausible range, then the rest of the parameters are identified. In the 2 applications we consider, we restrict δ to be a nonpositive value in $[\tau, 0]$ for some constant $\tau < 0$, because there is no evidence that the antiretroviral drug could increase the probability of HIV acquisition. In particular, $\delta = 0$ indexes the exclusion restriction; that is, there is no treatment effect among drug noncompliers.

Given a fixed δ , the estimation can be cast as a missing-covariate problem in the statistical literature. The potential outcome $D(1)$ for a placebo recipient is missing completely at random in Rubin's (23) sense. The standard likelihood-based method can be used for estimation, with incomplete data being taken expectation in the observed-data likelihood. When there is a case-control sample selected from the drug arm for drug assay, as in the iPrEx trial and the CAPRISA 004 trial, additional missing data are generated for persons who were not selected into the case-control sample. Assuming random dropout and missing visits, the measurements for $D(1)$ for these drug recipients are missing at random, since the missingness depends on case-control status. This constitutes so-called 2-phase sampling in the statistical literature (24, 25): In the first phase, HIV infection status and treatment assignment are observed for everyone; in the second phase, a proportion of participants are sampled to measure more expensive covariates, such as drug presence. Statistical methods have been extensively studied for estimating regression coefficients when some of the covariates are missing due to 2-phase sampling. When the covariate partially observed is discrete, one can use an expectation-maximization algorithm to estimate the parameters in a generalized linear model (26). The confidence intervals of the estimated effect can be computed by inverting the observed-data information matrix (27). In the Web Appendix (available at <http://aje.oxfordjournals.org/>), we describe the likelihood and the expectation-maximization algorithm for estimating ($\alpha, \beta_0, \beta_1, \beta_2$) when a case-control sample from the drug arm is selected for drug assay. A simulation study

is also presented in Web Table 1 to show the validity of the estimation procedure.

Note that a similar efficacy estimate among compliers has been studied in a vitamin supplement intervention (10), though not constructed under the potential outcomes framework. Sommer and Zeger (10) considered the simplest scenario, in which all participants in the active arm have data on compliance, and they derived the treatment effect among compliers using a method-of-moments type of estimator for the proportions of compliers and for the HIV infection probability among the compliers $\Pr(Y(0) = 1|D(1) = 0)$, also exploiting the exclusion restriction. This estimator is not a maximum likelihood estimator and therefore may not be statistically efficient. Our causal estimand is based on the potential outcomes framework, estimated using a maximum likelihood method for missing covariate data. Our method accommodates case-control sampling for drug assay and, in addition to obtaining the efficacy estimate, provides a general sensitivity model that includes the exclusion restriction as a special case.

RESULTS

In the iPrEx trial, 36 out of 1,251 antiretroviral drug recipients who were included in the modified ITT analysis had emergent HIV infection, while 64 out of 1,248 placebo recipients were HIV-positive (1). Blood samples were taken for seroconverters at the infection visit and the seronegative controls matched by visit. In the correlation analysis between drug detection and HIV infection in the antiretroviral arm, at least one of the study drugs was detected in 3 of 34 subjects with HIV infection and in 22 of 43 seronegative control subjects (1). In the CAPRISA 004 trial, among 445 women randomized to the gel arm, there were 38 seroconverters (2, 5). Among 444 women assigned to receive the placebo gel, 60 became infected during the trial. Samples were available from 34 seroconverters and from 301 women who remained uninfected (5). Three of 34 seroconverters and 79 of 301 uninfected women had a tenofovir concentration of 1,000 ng/mL or more.

The estimated odds ratios and confidence intervals are shown in Table 1. In the iPrEx trial, because HIV infection is a rare event, we can interpret the parameter estimates on the relative-risk scale. Specifically, the value of the parameter of interest, β_2 , suggests that among participants who would have a detectable level of drug were they

assigned to the drug arm, the relative risk of the treatment effect is 0.097 (95% confidence interval (CI): 0.027, 0.352), a 90.3% reduction in infection risk. This effect estimate is slightly attenuated compared with the odds ratio estimate in a correlation analysis (odds ratio = 0.093) if we simply compare the drug-detected participants (3/34) with the drug-undetected participants in the drug arm (22/43), though our estimate has a causal-effect interpretation. Our results for β_1 suggest that the baseline risk of infection for persons who would have a detectable drug level if assigned to the drug arm is not very different from the baseline risk of infection for those who would not have a detectable drug level if assigned to the drug arm (odds ratio = 0.955, 95% CI: 0.327, 2.784). In other words, persons who comply with the pill-taking regimen are not very different in terms of HIV infection risk from those who do not comply. Together with the plausible exclusion restriction, this explains why the naive estimate and the causal estimate are similar.

In the CAPRISA 004 trial, our analysis shows that the prevention effect on HIV infection among women with a tenofovir concentration of 1,000 ng/mL or more has the odds ratio 0.104 (95% CI: 0.024, 0.447; $P = 0.002$), corresponding to a risk reduction of 89.6%, whereas the analysis by Karim et al. (5), which compares this group with placebo recipients, yielded a 74% reduction in HIV incidence. Our results also suggest that there is some evidence of an increased baseline HIV risk among compliers compared with noncompliers, consistent with findings from applicator-based adherence data (2). The reason might be that women with fewer sex acts may not adhere well to the coitus-related dosing strategy.

Figure 1 shows treatment effect estimates among drug compliers for the two studies under a range of values for the sensitivity parameter δ with a magnitude up to $\log(0.75)$, a value representing approximately 25% efficacy in the drug noncomplier subgroup. With an increasing amount of residual protection in noncompliers, the point estimate of the efficacy among drug compliers decreases, though not substantially. This is understandable, since if there is residual protection in noncompliers, fewer HIV infections in the control arm will be attributable to the nonadherer group and correspondingly more HIV infections in the control arm will be attributable to the adherer group, which leads to a decrease in efficacy among compliers. Notably, the estimates in the iPrEx trial suggest strong efficacy even when $\delta = \log(0.75)$,

Table 1. Parameter Estimates for Regression Coefficients When $\delta = 0$ for the iPrEx Trial and the CAPRISA 004 Trial^a

	iPrEx Trial			CAPRISA 004 Trial		
	OR	95% CI	P Value	OR	95% CI	P Value
$\exp(\beta_0)$	0.055	0.035, 0.849	<0.001	0.115	0.081, 0.145	<0.001
$\exp(\beta_1)$	0.955	0.327, 2.784	0.933	2.606	0.801, 8.482	0.112
$\exp(\beta_2)$	0.097	0.027, 0.352	0.0004	0.104	0.024, 0.447	0.002

Abbreviations: AIDS, acquired immunodeficiency syndrome; CAPRISA, Centre for the AIDS Programme of Research in South Africa; CI, confidence interval; iPrEx, Preexposure Prophylaxis Initiative; OR, odds ratio.

^a Data were obtained from the iPrEx Study Team (1) and Karim et al. (2, 5) for the iPrEx Trial and the CAPRISA 004 Trial, respectively.

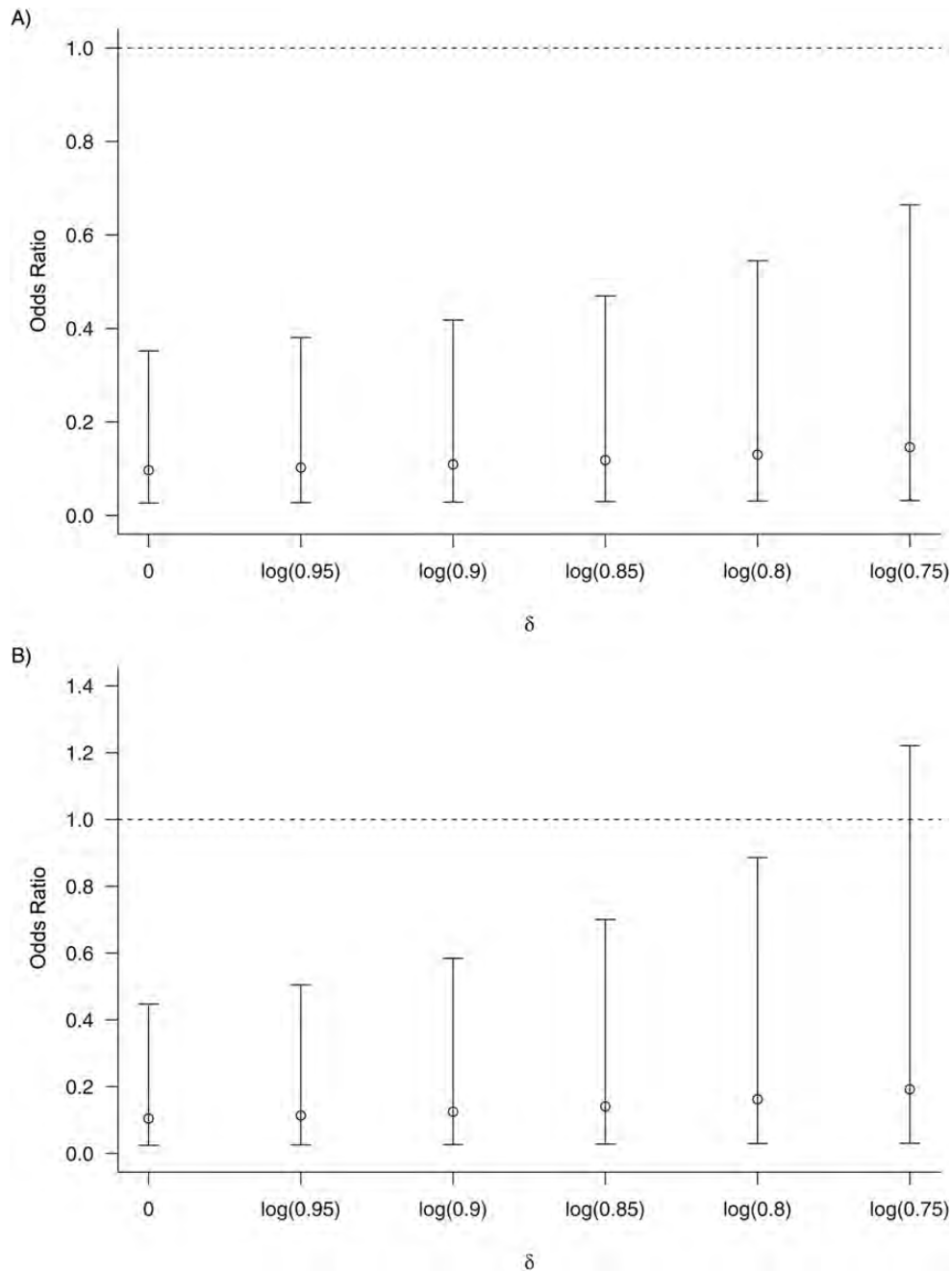


Figure 1. Sensitivity analysis for assessing the impact of deviating from the exclusion restriction on an estimate of efficacy among drug compliers, $\exp(\beta_2)$. Delta (δ) is the sensitivity parameter that indexes the degree of the deviation, with $\delta=0$ indicating the exclusion restriction. The upper panel shows the estimated treatment odds ratio among participants who have a detectable level of drug concentration in the Preexposure Prophylaxis Initiative (iPrEx) trial (1), and the lower panel shows the estimated treatment odds ratio among women who have a high drug concentration in the Centre for the AIDS Programme of Research in South Africa (CAPRISA) 004 trial (2, 5). Bars, 95% confidence interval. AIDS, acquired immunodeficiency syndrome.

while the confidence intervals in the CAPRISA 004 trial include 1 once δ approximates $\log(0.75)$. The wider confidence interval with a larger magnitude of δ , as seen in both trials, is due to the fact that fewer HIV infections are attributable to adherent placebo recipients, which leads to less stable treatment effect estimates among adherers.

DISCUSSION

We have defined and estimated PrEP efficacy among drug compliers using drug concentrations as a marker of adherence. Our estimation method applies regardless of whether compliance is equal across trial arms, since we

merely assume that drug noncompliers receive no protection from the treatment. Our analyses estimated a prevention efficacy of approximately 90% among drug compliers in the iPrEx trial and the CAPRISA 004 trial. These estimates are much greater than the primary ITT estimates for both trials; and the estimate for the iPrEx trial is fairly robust to deviation from the exclusion restriction in our sensitivity analysis, while the estimate in the CAPRISA 004 trial is less so. These results shed light on the importance of improving adherence in PrEP trials and in real-world applications once these products are licensed.

Our estimand was based on a binary classification of participants as drug compliers and drug noncompliers, which is probably an overly simplistic summary of a longitudinal and continuous spectrum of adherence. The estimand is essentially a subgroup effect, which limits its generalizability. With further strong and nonidentifiable assumptions, the population average treatment effect could be pursued to address the question, What is the average treatment effect if all participants are compliant? Examples of such assumptions include the average treatment effect among the treated being the same across the 2 trial arms and, within each arm, the average treatment effect for the treated being the same as that for the untreated, assuming hypothetically that all participants can be persuaded to adhere (28, 29).

Our analyses ignored the time-to-infection endpoint in both trials, merely modeling HIV infection as a binary indicator through logistic regression. The HIV hazard ratio in a proportional hazards model, rather than the odds ratio or risk ratio, is perhaps more appropriate for the risk-set sampling that actually took place in the iPrEx trial. For a rare event such as HIV infection, however, assuming that the baseline hazard of HIV infection is constant may be reasonable, so the regression estimates from a logistic model approximate the log hazard ratios in a proportional hazards model (30). Further work is under way to extend the estimation to failure-time outcomes in a nested case-control sampling scheme.

Other than invoking a sensitivity model, an alternative method of identifying the causal estimand is to build a model of $D(1)$ and some baseline characteristics in drug recipients and then predict $D(1)$ in the placebo recipients, similar to the methods of Follmann (11, 19). If blinding is truly well preserved, we can also use pill counts or gel applicator counts taken in the placebo recipients to predict their potential drug level if they were assigned to receive the active drug. A finer grid of compliance than simple dichotomization can be studied using this approach. These methods can provide estimates of the causal biological effects of PrEP drugs that will help to advance HIV prevention research.

ACKNOWLEDGMENTS

Author affiliations: Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington (James Y. Dai, Peter B. Gilbert, James P. Hughes, Elizabeth R. Brown); and Department of Biostatistics, School of Public Health, University of Washington, Seattle, Washington (James Y. Dai, Peter B. Gilbert, James P. Hughes, Elizabeth R. Brown).

This work was supported by National Institutes of Health grants U01 AI068615, 2 R37 AI054165-12, and R01 AI089341.

Conflict of interest: none declared.

REFERENCES

1. The iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *N Engl J Med*. 2010;363(27):2587–2599.
2. Karim QA, Karim SS, Frohlich JA, et al. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. *Science*. 2010; 329(5996):1168–1174.
3. Kelesidis T, Landovitz RJ. Preexposure prophylaxis for HIV prevention. *Curr HIV/AIDS Rep*. 2011;8(2): 94–103.
4. Skoler-Karpoft S, Ramjee G, Ahmed K, et al. Efficacy of Carraguard for prevention of HIV infection in women in South Africa: a randomised, double-blind, placebo-controlled trial. *Lancet*. 2008;372(9654):1977–1987.
5. Karim SS, Kashuba AD, Werner L, et al. Drug concentrations after topical and oral antiretroviral pre-exposure prophylaxis: implications for HIV prevention in women. *Lancet*. 2011; 378(9787):279–281.
6. Splawa-Neyman J. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Stat Sci*. 1990;5(4):465–472.
7. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol*. 1974;66(5):688–701.
8. Rubin DB. Bayesian inference for causal effects: the role of randomization. *Ann Stat*. 1978;6(1):34–58.
9. Holland PW. Statistics and causal inference. *J Am Stat Assoc*. 1986;81(396):945–960.
10. Sommer A, Zeger SL. On estimating efficacy from clinical trials. *Stat Med*. 1991;10(1):45–52.
11. Follmann DA. On the effect of treatment among would-be treatment compliers: an analysis of the Multiple Risk Factor Intervention Trial. *J Am Stat Assoc*. 2000;95(452): 1101–1109.
12. Rubin DR. More powerful randomization-based p -values in double-blind trials with non-compliance. *Stat Med*. 1998; 17(3):371–385.
13. Imbens GW, Rubin DB. Bayesian inference for causal effects in randomized experiments with noncompliance. *Ann Stat*. 1997;25(1):305–327.
14. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc*. 1996;91(434):444–455.
15. Frangakis CE, Rubin DB. Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika*. 1999;86(2):365–379.
16. Robins J, Rotnitzky A. Estimation of treatment effects in randomised trials with non-compliance and a dichotomous outcome using structural mean models. *Biometrika*. 2004; 91(4):763–783.
17. Rosenbaum PR. The consequences of adjustment for a concomitant variable that has been affected by treatment. *J R Stat Soc Ser A*. 1984;147(5):656–666.
18. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics*. 2002;58(1):21–29.

19. Follmann D. Augmented designs to assess immune response in vaccine trials. *Biometrics*. 2006;62(4):1161–1169.
20. Gilbert PB, Hudgens MG. Evaluating candidate principal surrogate endpoints. *Biometrics*. 2008;64(4):1146–1154.
21. Gilbert PB, Bosch R, Hudgens MG. Sensitivity analysis for the assessment of vaccine effects on viral load in HIV vaccine trials. *Biometrics*. 2003;59(3):531–541.
22. Shepherd BE, Gilbert PB, Dupont TC. Sensitivity analyses comparing time-to-event outcomes only existing in a subset selected post-randomization and relaxing monotonicity. *Biometrics*. 2011;67(3):1100–1110.
23. Rubin DB. Inference and missing data. *Biometrika*. 1976;63:581–592.
24. White JE. A two-stage design for the study of the relationship between a rare exposure and a rare disease. *Am J Epidemiol*. 1982;115(1):119–128.
25. Scott AJ, Wild CJ. Fitting logistic regression models in stratified case-control studies. *Biometrics*. 1991;47(2):497–510.
26. Ibrahim JG. Incomplete data in generalized linear models. *J Am Stat Assoc*. 1990;85(411):765–769.
27. Louis TA. Finding the observed information when using the EM algorithm. *J R Stat Soc Series B*. 1982;44(2):226–233.
28. Robins JM, Greenland S. Comment on Angrist, Imbens and Rubin: estimation of the global average treatment effects using instrumental variables. *J Am Stat Assoc*. 1996;91(434):456–458.
29. Robins JM. Correction for non-compliance in equivalence trials. *Stat Med*. 1997;17(3):269–302.
30. Green MS, Symons MJ. A comparison of the logistic risk function and the proportional hazards model in prospective epidemiologic studies. *J Chronic Dis*. 1983;36(10):715–724.