# AUGMENTED HIV VACCINE TRIAL DESIGN FOR ESTIMATING REDUCTION IN INFECTIOUSNESS AND PROTECTIVE EFFICACY

SUSMITA DATTA[†], M. ELIZABETH HALLORAN* AND IRA M. LONGINI, Jr.

*Department of Biostatistics, The Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta GA 30322, U.S.A.*

## SUMMARY

It is important to design HIV vaccine trials to estimate the efficacy of a vaccine in reducing infectiousness in addition to the protective efficacy. Currently planned phase III HIV vaccine field trials in which at-risk individuals are randomized and followed over time do not permit estimation or testing of the vaccine's effect on reducing infectiousness of vaccinees who become infected. We suggest an augmentation of these field trials that recruits steady sexual partners of the primary participants into the trial as far as they are willing to participate. This study design would allow estimation of the efficacy of the vaccine on reducing infectiousness as well as the protective efficacy. We compare the classical design that does not include partners to two different types of augmented design. In the first type of augmentation, called the *non-randomized partner design*, the steady sexual partners are not randomized to vaccine or placebo. In the second type of augmentation, called the *randomized partner design*, the steady sexual partners are also randomized to vaccine or placebo. We present a probability model based on infection status at the end of the trial that provides maximum likelihood estimates of the protective efficacy of the vaccine, $VE_S$, and the efficacy of the vaccine on reducing infectiousness, $VE_I$. Wald statistics are used for one degree of freedom tests on $VE_S$ and $VE_I$. With the augmented design, a likelihood ratio test is used to test whether the vaccine has any effect at all. The randomized partner design has more power and provides narrower confidence intervals than does the non-randomized partner design. © 1998 John Wiley & Sons, Ltd.

## 1. INTRODUCTION

Protective efficacy of a vaccine in reducing susceptibility to infection or disease is generally measured as $VE_S = 1 - RR$, where RR is the relative risk of the vaccinated compared to the unvaccinated group. The currently planned HIV vaccine field trials are designed to randomize individuals[1,2] to either vaccine or placebo and to base their estimates on (possibly interval-censored) time-to-infection data, and conditional on infection, progression to disease. The long-term and overall effects of a vaccination programme,[3] however, depend on much more than

* Correspondence to: M. E. Halloran, Department of Biostatistics, The Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322, U.S.A.

[†] Current Address: Department of Mathematics and Computer Science, Georgia State University, Atlanta, GA, U.S.A.

just protection against infection. They also depend on whether the vaccine reduces infectiousness of breakthrough infections, the duration of protection, whether the conferred protection is sensitive to boosting by exposure to natural infection,[4,5] and whether the behaviour of the vaccinees changes in response to vaccination.[6] If the immune response induced by an HIV vaccine does not protect well against infection, but reduces infectiousness if a person becomes infected, then the vaccine could be an important public health intervention. Thus, it is important to assess the effect of the vaccine on reducing infectiousness.

Vaccine trials based on relative risk parameters that do not condition on exposure to infection, such as events per person-time, hazards, or cumulative incidence, do not allow estimation of the transmission probability.[7] To measure the efficacy of a vaccine in reducing infectiousness of breakthrough infections, $VE_I$, we need to observe the relative risk of transmission from vaccinated infectives compared with unvaccinated infectives.[7,8] In HIV vaccine studies, this could be achieved by observing transmission within partnerships[9-14] with vaccine status as the covariate of interest. Koopman and Little,[15] using a special case of the model of Longini and Koopman,[16] suggest recruiting couples into a vaccine study, then randomizing by couple. The probability model includes the within-partnership secondary attack rate, allowing estimation of the effect of the vaccine on infectiousness from final value data. This approach would, however, preclude recruiting individuals with no steady sexual partners or with partners who do not wish to participate in the vaccine study. Rida[17] used a martingale approach to derive approximate confidence intervals for a similar design where just one partner is exposed to risk of infection outside the partnership.

We suggest a feasible augmentation of the currently planned phase III HIV vaccine field studies by including steady sexual partners of the primary randomized individuals as far as they are willing to participate.[18] This study design allows estimation of the vaccine efficacy for reducing infectiousness while preserving the classical vaccine trial design to allow estimation of protective efficacy. In the next section, we consider different aspects of these augmented study designs. In Section 3, we introduce a probability model and methods of estimation and hypothesis testing of the corresponding vaccine efficacy parameters that require just the vaccination and infection status of the primary participants and any partners included in the study. In Section 4, we present a simulation study to demonstrate how this augmented study design allows estimation of both protective efficacy and efficacy in reducing infectiousness and to demonstrate the statistical properties of the different study designs.

## 2. AUGMENTED TRIAL DESIGNS

Suppose we recruit a total of $n$ at-risk individuals into the study, with $n_1$ and $n_0$ randomized to either vaccine or placebo, respectively, with $n_1 + n_0 = n$. The outcome of interest is infection or progression to disease conditional on becoming infected, though we do not consider this latter case further here. Call these initially recruited individuals the primary participants. We refer to this design with randomized individuals only as the *classical* or *unconditional*[7] design. Without information on actual contact of the trial participants with potentially infectious persons, this trial design does not provide the information required to estimate the effect of the vaccine on reducing infectiousness.[19]

Suppose that some of the primary participants have susceptible steady sexual partners who are also willing to be recruited into the trial. Then there are an additional $n'_1$ and $n'_0$ partners of vaccinated and unvaccinated primary participants in the study, respectively. The recruited

partners could participate in several different ways. They may be willing to be randomized to vaccine or placebo as well. In other cases, they may simply be willing to have their infection status ascertained, but not to be randomized into the different arms. In this latter case, these people would respond biologically to a given infectious challenge the same as people randomized to placebo, although their behaviour, thus exposure to infection, might differ.

We could recruit steady partners at the beginning of or during the trial. The initially recruited steady sexual partnership could break up while others are formed during the trial. We need to accommodate the situation encountered in a particular trial in the final choice of probability model appropriate for the dynamics and data collected.

An advantage of choosing a design that augments the current design is that it allows recruitment of high risk individuals who do not have steady partners, but who can still contribute important information on the protective efficacy of the vaccine, $VE_S$. The recruitment of steady sexual partners of the primary participants, however, also allows estimation of the effect of the vaccine on reducing infectiousness, $VE_I$. The additional participants also provide additional information on the protective effects of the vaccine. Their exact contribution to this aspect depends on how we include them in the trial.

In the remainder of this paper, we consider two particular augmented design. In the first, called the *non-randomized partner design*, none of the recruited partners receives vaccine. In the second, called the *randomized partner design*, we randomize the recruited partners either to vaccine or to placebo.

## 3. A PROBABILITY MODEL AND INFERENCE FOR AUGMENTED DESIGNS

### 3.1. The probability model

Let $\theta$ denote the relative susceptibility of a vaccinated uninfected compared to an unvaccinated uninfected person, and $\phi$ denote the relative infectiousness of a vaccinated infected compared to an unvaccinated infected person. Then the desired estimates are the protective efficacy on susceptibility and the efficacy in reducing infectiousness, respectively,

$$VE_S = 1 - \theta, \quad VE_I = 1 - \phi.$$

Our goal is to find the maximum likelihood estimates of $\theta$ and $\phi$, and thus $VE_S$ and $VE_I$. For each individual, or individuals within partnerships, the required data are: (i) vaccine status; (ii) infection status at the end of trial; and (iii) vaccine and infection status of partner, if the study subject has a partner.

Let:

$v$ denote vaccination status of primary study participants, 0 if unvaccinated, 1 if vaccinated;
$\mu$ denote vaccination status of steady partners, 0 if unvaccinated, 1 if vaccinated;
$i, j$ denote infection status of primary participant and steady partners, respectively, where 0 denotes uninfected and 1 denotes infected by the end of the study;
$\gamma$ equal the probability of an unvaccinated trial participant becoming infected during the study period. For someone with a partner in the study, $\gamma$ represents the probability of being infected from sources outside the partnership;
$\beta_{v\mu}^{ij}$ equal the per partnership transmission probability of partnerships with vaccination status $(v, \mu)$, where either $i$ or $j$ equals 1.

## 3.2. Probability model for trial design without steady partners

The probability model is an extension of a model introduced by Longini and Koopman.[16] Consider the vaccine trial design based simply on enrolling at-risk individuals and randomizing them either to placebo or vaccine. Let $n$ denote the total number of primary participants in the trial, where $n_{0x}$ and $n_{1x}$ are the number of unvaccinated and vaccinated participants without partners in the study, respectively, that is, $n = n_{0x} + n_{1x}$. The $x$ denotes having no partner in the study. Since none of the participants in this study design has any steady sexual partner, $\gamma$ is the probability an unvaccinated participant becomes infected from any source over the course of the study. We assume that the probability has a vaccinated person becomes infected during the study is $\gamma_1 = \theta\gamma$, where $\theta$ is the relative susceptibility of the vaccinated compared to the unvaccinated. This represents an approximation since $\gamma$ is actually a function of the susceptibility of the unvaccinated individuals.[3,6]

Let $\pi_v^i$ be the probability that a primary participant with vaccination status $v$ has infection outcome $i$ over the course of the study. For unvaccinated primary participants, the probabilities of the possible outcomes are

$$\pi_0^0 = 1 - \gamma, \quad \pi_0^1 = \gamma. \tag{1}$$

For vaccinated primary participants, the probabilities of the possible outcomes are

$$\pi_1^0 = 1 - \theta\gamma, \quad \pi_1^1 = \theta\gamma. \tag{2}$$

## 3.3. Augmented designs

In the augmented designs, we recruit steady sexual partners of the primary participants into the study. Let $n'$ denote the total number of steady sexual partners recruited into the study, so that the total number of study participants is $N = n + n'$. If $n_0'$ and $n_1'$ denote the number of recruited steady sexual partners of unvaccinated and vaccinated primary participants, respectively, then $n' = n_0' + n_1'$. These recruited partners may or may not be randomized to either placebo or vaccine. The composition of the study participants is:

$n_{0x}$ = number of unvaccinated primary participants who lack steady sexual partners;
$n_{1x}$ = number of vaccinated primary participants who lack steady sexual partners;
$n_{00}$ = number of unvaccinated primary participants with unvaccinated steady sexual partners;
$n_{10}$ = number of vaccinated primary participants with unvaccinated steady sexual partners;
$n_{01}$ = number of unvaccinated primary participants with vaccinated steady sexual partners;
$n_{11}$ = number of vaccinated primary participants with vaccinated steady sexual partners.

Let $n_{v\mu}'$, $v = 0, 1$, $\mu = 0, 1$ be the corresponding number of steady partners augmenting the study in partnerships of vaccine status $(v, \mu)$. In this paper, we assume that each primary participant has at most one steady partner in the study, so that $n_{v\mu} = n_{v\mu}'$. This condition is not necessary for the model. It is quite likely in field trials that some participants would have more than one steady partner, or change their partners over time. We could easily incorporate these situations into the model.[16]

The composition of the study participants in the non-randomized partner design is

$$n_v = n_{vx} + n_{v0}, \quad n_v' = n_{v0}', \quad v = 0, 1.$$

In the randomized partner design, the composition is

$$n_v = n_{vx} + n_{v0} + n_{v1}, \quad n'_v = n'_{v0} + n'_{v1}, \qquad v = 0, 1.$$

The probability model assumes that (i) primary participants and recruited steady partners are equally at risk for infection outside the partnership, (ii) primary participants and recruited steady partners are equally susceptible and equally infectious if infected, and (iii) vaccine efficacy is the same in primary participants and in recruited partners. Under assumption (ii), we denote $\beta_{00}^{01} = \beta_{00}^{10} = \beta$, the transmission probability of infection within the partnership when both partners are unvaccinated. Similarly, $\beta_{11}^{01} = \beta_{11}^{10}$ denotes the transmission probability when both partners are vaccinated. These assumptions might not hold, for instance, in vaccine trials with heterosexual partnership, where there is the belief that the transmission probability from women to men is lower than that from men to women. We would use the more general model for this situation.

The relative infectiousness is the ratio of the probability of transmission from a vaccinated infected person to his or her unvaccinated susceptible partner compared to the probability of transmission if both partners are unvaccinated:

$$\phi = \frac{\beta_{10}^{10}}{\beta} = \frac{\beta_{01}^{01}}{\beta}.$$

For simplicity, we assume that the protective effects of the vaccine in reducing the probability of infection over the course of the study equals the protective effect of the vaccine in reducing transmission from an unvaccinated infective within a partnership:[15]

$$\theta = \frac{\gamma_1}{\gamma} = \frac{\beta_{01}^{10}}{\beta} = \frac{\beta_{10}^{01}}{\beta}.$$

Furthermore, we assume that within a partnership in which both are vaccinated, the reduction in infectiousness and the reduction in susceptibility due to the vaccine are independent and do not interact, so that

$$\phi\theta = \frac{\beta_{11}^{01}}{\beta} = \frac{\beta_{11}^{10}}{\beta}.$$

The probabilities $\pi_v^i$ of the primary participants without steady partners in the study becoming infected during the period of observation are the same as in equations (1) and (2).

For partnerships in the study, let $\pi_{v\mu}^{ij}$ be the probability that the primary participant and steady sexual partner with vaccine status $v$ and $\mu$, respectively, have the infection outcomes $i$ and $j$ at the end of the study, respectively. When both partners are unvaccinated, the probabilities of the possible outcomes are

$$\pi_{00}^{00} = (1 - \gamma)^2$$

$$\pi_{00}^{10} = \pi_{00}^{01} = \gamma(1 - \beta)(1 - \gamma)$$

$$\pi_{00}^{11} = \gamma^2 + 2\gamma(1 - \gamma)\beta.$$

When one of the partners is vaccinated and the other is unvaccinated, the probabilities of the possible outcomes are

$$\pi_{01}^{00} = \pi_{10}^{00} = (1 - \theta\gamma)(1 - \gamma)$$

$$\pi_{01}^{10} = \pi_{10}^{01} = \gamma(1 - \theta\gamma)(1 - \theta\beta)$$

$$\pi_{01}^{01} = \pi_{10}^{10} = (1 - \gamma)\theta\gamma\,(1 - \phi\beta)$$

$$\pi_{01}^{11} = \pi_{10}^{11} = \theta\gamma^2 + \gamma(1 - \theta\gamma)\theta\beta + \theta\gamma(1 - \gamma)\phi\beta.$$

When both partners are vaccinated, the probabilities of the possible outcomes are

$$\pi_{11}^{00} = (1 - \theta\gamma)^2$$

$$\pi_{11}^{01} = \pi_{11}^{10} = \theta\gamma(1 - \theta\gamma)(1 - \theta\phi\beta)$$

$$\pi_{11}^{11} = (\theta\gamma)^2 + 2\theta\gamma(1 - \theta\gamma)\,\phi\theta\beta. \tag{3}$$

In the Appendix, we consider extension of the model to account for multiple partners. Since the probability model includes parameters for the risk of infection during the study period and the probability of becoming infected upon exposure to an infected person, the model combines two levels of analysis from the hierarchy of parameters defined in Rhodes *et al.*[20] The transmission probability, or the steady sexual partner secondary attack rate in this case, is a level I parameter that conditions on contacts between susceptibles and infectives. The cumulative incidence of infection is an unconditional level IV parameter that requires only information on whether people become infected during the course of the study. When used to estimate protective efficacy or effectiveness of a vaccine, these parameters all require equal exposure to infection in the vaccinated and unvaccinated groups.[21] The assumption that we can use the same $\theta$ parameter to estimate protective efficacy for the two levels is an approximation.[3,6] We could establish a probability model to include two different parameters. In cases of low incidence and assumptions of homogeneous exposure and susceptibility,[22,23] the approximation is appropriate.

### 3.4. Inference for designs without partnerships

Let $m_v^i$ be the observed frequency for the $i$th, $i = 0, 1$, outcome with primary participants of vaccine status $v$, $v = 0, 1$ of the $\pi_v^i$ given in equations (1) and (2). The likelihood function for estimating $\gamma$ and $\theta$ is

$$L(\gamma, \theta) \propto \prod_{i=0}^{1} \prod_{v=0}^{1} (\pi_v^i)^{m_v^i}. \tag{4}$$

We let $AR_0$ and $AR_1$ be the cumulative incidences of infection (attack rates, AR) in the unvaccinated and vaccinated groups, respectively. By taking the partial derivatives of the log-likelihood function and equating them to zero we get the maximum likelihood estimates of $\gamma$ and $\theta$:

$$\hat{\gamma} = \frac{m_0^1}{m_0^0 + m_0^1} = AR_0, \quad \hat{\theta} = \frac{m_1^1}{(m_1^0 + m_1^1)\hat{\gamma}} = AR_1/AR_0.$$

From the estimate of $\theta$, we can find the estimate of vaccine efficacy, $\widehat{VE}_S = 1 - \hat{\theta}$. We use the asymptotic standard deviation to construct a confidence interval on $VE_S$. The one degree of freedom hypothesis test of a protective effect of the vaccine, $VE_S$, is

$$H_0 : VE_S = 0 \text{ versus } H_1 : VE_S \neq 0.$$

We perform the hypothesis test using the large-sample Wald statistic

$$Z_{VE_S} = \frac{\widehat{VE}_S}{\text{std}(\widehat{VE}_S)}$$

to test whether $|Z_{VE_S}| > 1\cdot96$ or not, where $\text{std}(\widehat{VE}_S)$ denotes the standard deviation of $\widehat{VE}_S$. We emphasize that in this model, $VE_I$, the effect of the vaccine on infectiousness, is unestimable.

## 3.5. Inference for augmented designs

When we augment the design with inclusion of partners in the study, let $m_{v\mu}^{ij}$ denote the observed frequency of the $i, j$ outcome in partnerships of vaccine status $v, \mu$ at the end of the study. We construct the likelihood function to estimate $\beta$ and $\phi$ as well as $\gamma$ and $\theta$:

$$L(\gamma, \beta, \theta, \phi) \propto \prod_{i=0}^{1} \prod_{v=0}^{1} (\pi_v^i)^{m_v^i} \prod_{j=0}^{1} \prod_{\mu=0}^{1} (\pi_{v\mu}^{ij})^{m_{v\mu}^{ij}}. \tag{5}$$

If none of the steady sexual partners recruited into the study is vaccinated, then $\mu$, the vaccination status of the partner, is always zero, while for the randomized partner design, $\mu = 0, 1$. The likelihood function for the non-randomized partner design is

$$L(\gamma, \beta, \theta, \phi) \propto \prod_{i=0}^{1} \prod_{v=0}^{1} (\pi_v^i)^{m_v^i} \prod_{j=0}^{1} (\pi_{v0}^{ij})^{m_{v0}^{ij}}. \tag{6}$$

We obtain the maximum likelihood estimates of the parameters in the model by simultaneously solving the score equations

$$\frac{\partial \ln L}{\partial \gamma} = 0, \quad \frac{\partial \ln L}{\partial \beta} = 0, \quad \frac{\partial \ln L}{\partial \theta} = 0, \quad \frac{\partial \ln L}{\partial \phi} = 0.$$

In practice, we seek numerical solutions using the Newton–Raphson method starting from some initial method of moment type estimates of the parameters. We can in turn estimate the vaccine efficacies $VE_S$ and $VE_I$ by $\widehat{VE}_S = 1 - \hat{\theta}$, $\widehat{VE}_I = 1 - \hat{\phi}$. Usually we have interest in testing the following hypotheses:

$$H_0 : VE_S = 0 \text{ versus } H_1 : VE_S \neq 0,$$

$$H_0 : VE_I = 0 \text{ versus } H_1 : VE_I \neq 0.$$

The first hypothesis tests whether the vaccine has an effect on susceptibility, whereas the second one tests whether the vaccine has an effect on infectiousness. We can form large-sample Wald statistics as in the previous section.

In this model, however, it is also possible to carry out a simultaneous test of both types of vaccine efficacies. The appropriate null hypothesis for this problem

$$H_0: \text{VE}_\text{S} = \text{VE}_\text{I} = 0$$

$$H_1: \text{at least one differs from zero.}$$

The hypothesis tests whether the vaccine has any effect at all, either on susceptibility or infectiousness, or not. An appropriate test for this problem is the likelihood ratio test. Under the null hypothesis, the likelihood is a function of $\gamma$ and $\beta$ only and obtains from the full likelihood by substitution $\theta = \phi = 1$. This test should have better power than the individual tests for $\text{VE}_\text{S}$ and $\text{VE}_\text{I}$ mentioned earlier.

## 4. SIMULATION STUDIES

### 4.1. Methods of simulation

We conducted simulation studies to investigate the sampling distribution characteristics of the estimators of the vaccine efficacy parameters for the three different models specified in the earlier section. Using Monte Carlo simulations, we calculated the bias, variance and the mean square error of the estimates of $\text{VE}_\text{S}$ and $\text{VE}_\text{I}$. We also studied the power of hypothesis tests with the efficacy parameters $\text{VE}_\text{S}$ and $\text{VE}_\text{I}$. For the classical trial, we generated 500 Monte Carlo simulations of total sample size 4000 each. For the simulation we assume the same number of participants in the vaccinated and unvaccinated groups ($n_0 = n_1 = 2000$). The probability that an unvaccinated person becomes infected during the study is $\gamma = 0.1$, and the proportion by which the vaccine reduces the susceptibility of the vaccinated individual is $\theta = 0.8$, that is, $\text{VE}_\text{S} = 0.2$. From each sample, we estimated $\text{VE}_\text{S}$.

We simulated two augmented study designs at two levels of augmentation. In each case, there were $n = 4000$ primary participants with 2000 randomized either to vaccine or placebo. In the non-randomized partner design, none of the recruited steady sexual partners receive vaccine. At the first simulated level of augmentation, 500 primary participants in each vaccine trial arm recruit steady sexual partners into the study:

$$n_{0x} = n_{1x} = 1500, \quad n_{00} = n'_{00} = n_{10} = n'_{10} = 500.$$

At the second level of augmentation, 1000 primary participants in each vaccine trial arm recruit steady sexual partners into the study:

$$n_{0x} = n_{1x} = 1000 = n_{00} = n'_{00} = 1000, \quad n_{10} = n'_{10} = 1000.$$

In the randomized partner design, we assumed that half of the recruited steady partners in each arm receive placebo and half receive vaccine. The composition of the 4000 primary participants for the two levels of augmentation are, respectively

$$n_{0x} = n_{1x} = 1500, \quad n_{00} = n_{01} = n_{10} = n_{11} = n'_{00} = n'_{01} = n'_{10} = n'_{11} = 250.$$

and

$$n_{0x} = n_{1x} = 1000, \quad n_{00} = n_{01} = n_{10} = n_{11} = n'_{00} = n'_{01} = n'_{10} = n'_{11} = 500.$$

Table I. Estimation and hypothesis testing results for the classical trial and the two augmented designs with 500 steady sexual partners per arm. Parameters used for simulation are $\gamma = 0\cdot1$, $\theta = 0\cdot8$, $\phi = 0\cdot5$ and $\beta = 0\cdot5$

|  | Parameters of interest | Classical trial | Non-randomized partner model | Randomized partner model |
|---|---|---|---|---|
| Mean | $VE_S$ | 0·198 | 0·198 | 0·199 |
|  | $VE_I$ |  | 0·501 | 0·484 |
| Standard deviation | $VE_S$ | 0·082 | 0·072 | 0·067 |
|  | $VE_I$ |  | 0·057 | 0·026 |
| Mean squared error | $VE_S$ | 0·007 | 0·005 | 0·004 |
|  | $VE_I$ |  | 0·057 | 0·026 |
| Confidence interval | $VE_S$ | 0·022–0·345 | 0·049–0·317 | 0·057–0·323 |
|  | $VE_I$ |  | 0·005–0·952 | 0·111–0·778 |
| Power | $VE_S$ | 0·668 | 0·784 | 0·846 |
|  | $VE_I$ |  | 0·570 | 0·864 |
| Power of the LR test | $VE_S$ and $VE_I$ |  | 0·856 | 0·964 |

Table II. Estimation and hypothesis testing results for the classical trial and the two augmented designs with 1000 steady sexual partners per arm. Parameters used for simulation are $\gamma = 0\cdot1$, $\theta = 0\cdot8$, $\phi = 0\cdot5$ and $\beta = 0\cdot5$

|  | Parameters of interest | Classical trial | Non-randomized partner model | Randomized partner model |
|---|---|---|---|---|
| Mean | $VE_S$ | 0·198 | 0·196 | 0·198 |
|  | $VE_I$ |  | 0·502 | 0·499 |
| Standard deviation | $VE_S$ | 0·082 | 0·066 | 0·057 |
|  | $VE_I$ |  | 0·163 | 0·116 |
| Mean squared error | $VE_S$ | 0·007 | 0·004 | 0·003 |
|  | $VE_I$ |  | 0·027 | 0·013 |
| Confidence interval | $VE_S$ | 0·022–0·345 | 0·067–0·317 | 0·082–0·306 |
|  | $VE_I$ |  | 0·176–0·808 | 0·259–0·70 |
| Power | $VE_S$ | 0·668 | 0·822 | 0·932 |
|  | $VE_I$ |  | 0·870 | 0·988 |
| Power of the LR test | $VE_S$ and $VE_I$ |  | 0·966 | 0·996 |

We also examined the statistical properties of the estimators, holding fixed the total number of trial participants. Parameters used in the simulations of augmented trials designs in Tables I and II are $\gamma = 0\cdot1$, $\theta = 0\cdot8$, $\phi = 0\cdot5$, $\beta = 0\cdot5$. In Table III, we assume that the per partnership transmission probability is only half that in Tables I and II, $\beta = 0\cdot25$.

For each of the trial designs, we report the average of 500 estimates of the estimable efficacy parameters are their standard deviation. In the classical design, this is $VE_S$, and in the augmented designs this is $VE_S$ and $VE_I$. The average of the 500 estimates leads to an estimate of the bias. The standard deviation of the simulated estimates is the estimated standard deviation of the efficacy parameter estimate. We also report the equitailed interval ranging from the lower 2·5th to upper

Table III. Estimation and hypothesis testing results for the classical trial and the two augmented designs with 1000 steady sexual partners per arm. Parameters used for simulation are $\gamma = 0.1$, $\theta = 0.8$, $\phi = 0.5$ and $\beta = 0.25$

|  | Parameters of interest | Classical trial | Non-randomized partner model | Randomized partner model |
|---|---|---|---|---|
| Mean | $VE_S$ | 0·198 | 0·196 | 0·199 |
|  | $VE_I$ |  | 0·500 | 0·491 |
| Standard deviation | $VE_S$ | 0·082 | 0·072 | 0·064 |
|  | $VE_I$ |  | 0·276 | 0·206 |
| Mean squared error | $VE_S$ | 0·007 | 0·005 | 0·004 |
|  | $VE_I$ |  | 0·076 | 0·042 |
| Confidence interval | $VE_S$ | 0·022–0·345 | 0·038–0·325 | 0·052–0·314 |
|  | $VE_I$ |  | 0·000–0·999 | 0·000–0·887 |
| Power | $VE_S$ | 0·668 | 0·782 | 0·876 |
|  | $VE_I$ |  | 0·464 | 0·670 |
| Power of the LR test | $VE_S$ and $VE_I$ |  | 0·858 | 0·946 |

2·5th percentile of the estimated sampling distribution, an interval that we can regard as a 95 per cent confidence interval for the parameter estimate. To compare the overall performance of the efficacy parameter estimators under different probability models, we also report the mean square errors of the estimates. Lastly, we report the approximate power of the asymptotic hypothesis test at 5 per cent level $H_0$: $VE_S = 0$, versus $H_1$: $VE_S \neq 0$. The power is approximated by the proportion of times in the simulation $|\widehat{VE}_S/\text{std}(\widehat{VE}_S)|$ exceeds 1·96 (out of 500 simulation runs). This is justified because, under the null hypothesis, $\widehat{VE}_S/\text{std}(\widehat{VE}_S)$ is asymptotically standard normal.

Similarly, in the augmented designs, the approximate power of the 5 per cent test for $H_0$: $VE_I = 0$, versus $H_1$: $VE_I \neq 0$ is approximated by the proportion of times (out of 500 iteration) in the simulation where $|\widehat{VE}_I/\text{std}(\widehat{VE}_I)|$ exceeds 1·96. Lastly, to enumerate the power of the likelihood ratio test, we perform each of the simulations twice, once under the null and once under the alternative. We generate the samples under the null hypothesis, that is, we assume both $VE_S$ and $VE_I$ are zero and calculate the upper 5 per cent percentile of the null distribution of the test statistic. Then we generate samples from the alternative model, where the parameter values are as described in the beginning of this section. The power of the likelihood ratio test is approximated by the proportion of times the likelihood ratio test statistic in the simulation under the alternative model exceeds the upper 5 per cent percentile point of the null distribution calculated from a separate simulation mentioned before.

## 4.2. Simulation results

Tables I and II contain the simulation results. The second columns of the tables contain the results for the classical trial where we assume that principal participants are recruited without steady sexual partners. The results for the non-randomized partner model and randomized partner model appear in the third and fourth columns of the tables, respectively. First we discuss results in Table I, where we have information about 500 steady sexual partners for each group of

the vaccinated and unvaccinated principal participants. $VE_S = 0.2$ in all the simulations. The bias of the estimate of the vaccine efficacy parameter $VE_S$ is similar in all the models we considered, although there is a minor improvement of the bias in the randomized partner model. The standard deviation, however, is reduced in non-randomized model compared to the classical model, and even further reduced in the randomized partner model. The mean squared error (MSE) of estimation is smaller in the non-randomized model than in the classical model and is the smallest in the randomized partner model. The 95 per cent confidence intervals of the parameter estimates appear in the fourth row. We see a slight improvement, that is, narrower confidence intervals for $VE_S$ with the two augmented trials compared with the classical trial.

The power of the test of significance of the null hypothesis $H_0$: $VE_S = 0$ improve considerably in the two augmentations of the classical model compared with the classical model. In the classical model, the power is 0·668, in the non-randomized partner model it is 0·784, and in the randomized partner model it is 0·846. This increase in power is due partly to the increased number of participants, and partly to our taking into account the within-partnership transmission probability.

We simulated the power to reject $H_0$: $VE_S = 0$ for the classical trial design with an equal number of participants as the non-randomized design with 500 partners recruited among participants in each arm of the study, that is, $n_0 = 3000$ and $n_1 = 2000$. This power is 0·768 (not shown in Table I). If we conduct a trial using the classical design in an equivalently sized population as the randomized partner design, that is, $n_0 = 2500$ and $n_1 = 2500$, the simulated power is 0·786. Both of these simulations assume that the probability of being infected during the study is 0·1, and partnerships are not taken into account. Essentially the power of the study for $VE_S$ is similar for the classical and the augmented designs if the sample sizes are the same and we use the same underlying transmission assumptions to produce the number of events in the study.

Aside from the power considerations, we emphasize that in the classical model, it is not possible to estimate or to test for $VE_I$ at all. In contrast, we can estimate $VE_I$ with either of the augmented trial designs. Hence, we can compare the performance of the non-randomized partner model and randomized partner model solely with respect to the estimation and tests relating $VE_I$. While the bias is almost the same in the non-randomized and randomized partner designs, the standard deviation improves in the randomized partner model from 0·057 to 0·026. Overall, the mean squared error for the estimation is also smaller in the randomized partner model. The confidence intervals of the estimates of $VE_I$ is narrower in the randomized partner model compared to the non-randomized partner model. The power of the test involving the parameter $VE_I$ increases dramatically from 0·57 with the non-randomized model to 0·86 with the randomized partner model. This suggests that if we can augment the trial design by recruiting partners, there is considerable benefit to assign some of them to vaccine. The power of the likelihood ratio test of the joint hypothesis involving both $VE_I$ and $VE_S$ is also larger in the randomized partner model.

We provide similar comparisons of these three different models when 1000 principal participants in each group (vaccinated and unvaccinated) have steady sexual partners (Table II). The second column of Table II contains the same results of the classical trial with $n_0 = n_1 = 2000$ as Table I. Once again, the bias of the estimates of $VE_S$ remains almost the same in all the models. The standard deviations and the mean squared error follow the same pattern as before, namely, they are the largest for the classical trial and the smallest for the randomized partner

trial. We see the same patterns with the confidence intervals as well, that is, they are narrower with the augmented models than with the classical trial model, and narrower with the randomized partner model than with non-randomized partner model. The power of the test for $H_0$: $VE_S = 0$ also has the same ordering of the three models as before, for the reasons discussed above. The power of the test is 0·668 in the classical trial, 0·822 in the non-randomized partner model, and 0·932 in the randomized partner model. Comparison of the estimation and tests of $VE_I$ also reveals the same pattern. Finally, the power of the likelihood ratio test of the joint parameters improves from 0·966 with the non-randomized partner model to 0·996 with the randomized partner model.

We can also compare the performances of the tests as the number of steady sexual partners increases from 500 to 1000 per trial arm. With more partners recruited, we have more information. We expect that this extra information on the partners will result in better estimation and increased power for tests of $VE_I$. For example, in the non-randomized partner model, the power of the test involving $VE_I$ in Table I is 0·570, and the power of the same test in Table II is 0·870. Similarly, the power of the likelihood ratio test for the non-randomized partner model increases from 0·856 to 0·966 with the increase from 500 to 1000 steady sexual partners per study arm. The same trend holds for comparisons of the other statistical measures as the number of steady sexual partners added to the study increases.

In additional simulations with the same parameter values and study composition except that $VE_S = 0·5$, that is, $VE_S$ is greater than in the previous simulations, the power to estimate $VE_S > 0·999$ in each of the three study designs. For estimation of $VE_I$, however, the situation became worse. This is likely due to the reduced number of infected vaccinees exposing their partners to infection.[7] The complex interaction between $VE_S$, $VE_I$ and the requirements for sample size determination deserves further investigation.

In Table III, $\beta = 0·25$, just half of that in Tables I and II. The confidence interval on $VE_I$ is much wider and the power of the hypothesis tests for $VE_I$ are much lower at this lower transmission probability than that in Table II. The power for $VE_S$ is also lower because there are fewer total events with the lower transmission probability. If the assumed transmission probability is $\beta = 0·25$ at a particular trial site, then this sample size is inadequate for studying $VE_I$ at this assumed level of efficacy. Longini et al.[18] give asymptotic results for the size of approximate confidence intervals as functions of $\beta$ and $\gamma$.

Recall that in the classical trial model, the assumption is that none of the principal participants has a steady sexual partner. A more realistic situation is that some of the primary participants have steady partners who are not included in the study. For these primary participants, the simplest approach is to model their probability of infection over the course of the trial independent of the source of infection. To see how the estimator and the tests perform, in such a setup in which participants actually have partners, but they are not included in the study or taken into account in the analysis, we simulate the data from a (larger) non-random partner model, but use only the information (vaccine and infection status) on the primary participant and analyse the data as if they had come from a classical trial design. Statistical inference is then performed using the classical model likelihood. The MSE of the estimators in this modified classical trial is nearly identical to the classical trial. The power of the test $H_0$: $VE_S = 0$ in the modified classical trial where data are generated from a non-randomized partner model with $n_{0x} = n_{1x} = 1000$, $n_{00} = n_{10} = 1000$, turned out to be 0·676 compared to 0·668 for the classical trial with the same total sample size. The power is still lower than the other two augmented models for the same test.

## 5. DISCUSSION

We suggest an augmented design of phase III HIV vaccine trials that enables estimation of the effect of the vaccine on reducing infectiousness in addition to estimation of protective efficacy. Using a simple probability model, we have demonstrated that estimation of both the protective efficacy, $VE_S$, and the efficacy in reducing infectiousness, $VE_I$, is possible with this design. Indeed, the simulation results suggest that the extra information provided by inclusion of the steady sexual partners not only enables us to estimate the vaccine efficacy due to reduction of the infectiousness of the vaccinated infected individual $VE_I$ along with $VE_S$, it also improves the accuracy of the estimator of $VE_S$ and the power of the test related to $VE_S$, compared to the classical model. If the HIV vaccine trials are designed to estimate efficacy simply by the relative time to infection, or, conditional on infection, the rate of progression to disease, then they are unable to give any information on the effect of the vaccine on reducing infectiousness. The increased power and better estimation provided by the randomized partner design compared with the non-randomized partner design suggest that if we have the ability to augment the trial design by recruiting partners, we gain considerably with our further ability to assign some of them to vaccine.

The probability model we presented here is a simple one that has been used with other infectious diseases where final value data were available for households within a community.[16] Magder and Brookmeyer use a similar model for HIV transmission for studying the effect of covariates on the partnership secondary attack rates with partner studies.[24]

Rida[17] uses a counting process approach similar to that in Rhodes *et al.*[20] The transmission probability is defined as the per contact probability of transmission from an infected person to a susceptible person. Once the partner is infected, the probability that the contact is potentially infectious as defined in Rhodes *et al.*[20] is 1. The hazard rate of infection, say $h(t)$, is then the rate of contacts times the transmission probability. In the model used by Rida,[17] the risk of secondary infection is defined as the competing risk between the within-couple hazard of infection and the hazard that the index case recovers either by the natural loss of infection or by change in the risk behaviour whereby he or she no longer exposes the partner to HIV. Let $t_i$ denote the time of infection of the first partner, and $T_i$ denote either the infection time of the second partner, the end of the study, the end of infectiousness of the first infected partner, or the time at which the partners cease within-partnership risky behaviour. In the probability model in this paper, the secondary attack rate for any partnership $i$ is approximately equivalent to

$$\beta_i = 1 - \exp\left(-\int_{t_i}^{T_i} h(t)\, \mathrm{d}t\right)$$

in the notation used by Rida. The different cases with differing lengths of exposure of the uninfected partner are not distinguished in this simple mode, but subsumed in the use of the final value data. Thus, $\beta$ represents a population average for the secondary attack rate over the entire period of the study in this paper. The vaccine efficacy parameters $\theta$ and $\phi$ are defined in this paper multiplicative to $\beta$, the probability of being infected within partnership during the trial, whereas in Rida's paper relative to the per contact transmission probability, thus multiplicative to $h(t)$. Thus, the relation between the efficacy parameters in our paper and those in Rida's is similar to that between those based on the transmission probability, hazards, and final value data discussed elsewhere.[3, 6, 25]

The model of vaccine effect assumes a multiplicative or leaky[4] effect of the vaccine on both susceptibility and infectiousness. Other models for heterogeneous distribution of the vaccine effect on susceptibility have been discussed elsewhere (see, for example, references 3, 6, 22, 23, 25, 26 and 27). The distribution of the effect of the vaccine on infectiousness does not have to be the same as that on susceptibility. For example, if a vaccine has an all-or-none effect on susceptibility, that is, some people are completely protected against infection while others are not at all, the vaccine could still reduce infectiousness in those people not protected against infection. In addition, the probability of infection from outside contacts over the course of the study may be different for an individual who does not have a steady partner than for an individual who does. An individual without a steady partner may engage in more risky behaviour. Also, there could be host factors such as age or history of STDs that differ among individuals with and without steady partners that affect the susceptibility of infection. These relations deserve further investigation.

Longini, Halloran and co-workers have expanded the model to account for interval censored, time-to-infection data of the sort expected in these trials with the transmission probability defined per sex act, not per partnership, and to account for more complicated partnership formations (unpublished). We can, however, analyse time-to-event data in trials augmented with recruitment of steady sexual partners using the model presented here. The more complicated model of partnership formation can accommodate local social and sexual practices where trials are being planned. These differ considerably within the United States, and across continents where trials are planned from Brazil to Uganda to Thailand.[28] If we expect that many couples will cease risky behaviour if one partner becomes infected, then we may need to increase either the sample size or the duration of follow-up to maintain power, depending on the biologic assumptions regarding distribution of infectiousness after infection. Although the model solved here assumes that all the recruited partners are susceptible, we could expand the design to include partners who are already infected.

Previous trials on many different vaccines have often missed the opportunity to estimate the effect of the vaccine on reducing infectiousness of breakthrough cases because of inadequate design to estimate anything other than the relative risk of infection or disease in the vaccinated compared to the unvaccinated. An example of this is the recently licensed chickenpox vaccine in the U.S.A.[5] It is important to design vaccine trials to answer the important questions. We suggest augmentation of HIV vaccine trials with steady sexual partners to enable estimation of both the effect of the vaccine on infectiousness in addition to its protective effect.

## APPENDIX: LIKELIHOOD WHEN PRIMARY PARTICIPANT HAS MULTIPLE PARTNERS

Here we present the likelihood function for primary participants who have more than one steady partner. Suppose a primary participant has $r$ partners who are not randomized into the trial and that we can ascertain infection status of all or a subset of his or her sexual partners. We let array $\boldsymbol{v} = (v, \mu_1, \ldots, \mu_r)$ be the vaccination status indicators of the primary participant and his or her $r$ partners, and array $\mathbf{i} = (i, j_1, \ldots, j_r)$ be the corresponding infection indicators. Then we let $\pi_v^{\mathbf{i}}$ be the probability of outcome $\mathbf{i}$, for vaccination status $\boldsymbol{v}$. The particular form of $\pi_v^{\mathbf{i}}$ will depend on the structure of the sexual contacts among the $r + 1$ persons. The simplest structure is to assume that the sexual contacts between the primary participant and a particular sexual partner are independent of such contact with any other sexual partner. The infection status of the $r$ sexual partners

may be obtained through retrospective ascertainment of partners of infected primary participants, that is, $\mathbf{i} = (1, j_1, \ldots, j_r)$. The simplest assumption is that the contacts do not have an opportunity to infect each other. An additional simple assumption is that none of the sexual partners of primary participants will be vaccinated, that is, $\mathbf{v} = (v, 0, \ldots, 0)$. This corresponds to the non-randomized partner design.

Let $j = \sum_{k=1}^{r} j_k$ be the total number of partners infected. We assume that we can ascertain the infection status of all $r$ partners. Thus, one of the partners must be the source of infection for the infection primary participant. Then, the probability of interest is

$$\pi_v^{\mathbf{i}} = \binom{r-1}{j-1} [1 - (1-\gamma)(1-\phi^v \beta)]^{j-1} [(1-\gamma)(1-\phi^v \beta)]^{r-j}, \quad j = 1, \ldots, r, \quad v = 0, 1. \quad (7)$$

This probability can be factored into the appropriate likelihood function. If all of the $r$ partners cannot be located and tested, then (7) can be modified to allow that the source of infection for the primary participants may not be among the tested partners. This can be done using straightforward probability arguments. If some of the $r$ partners are known to have sexual contact with one another, then (7) can be formulated for that particular situation following the recursive methods developed by Addy *et al.*[29] We can expand expression (7) so that some of the partners are vaccinated. We can also account for the possible ordering of the partnerships. This leads to the introduction of time and other models being developed elsewhere.

## REFERENCES

1. Dixon, D. O., Rida, W. N., Fast, P. E. and Hoth, D. F. 'HIV vaccine trials: some design issues including sampling size calculations', *Journal of the Acquired Immune Deficiency Syndrome*, **6**, 485–496 (1993).
2. Fleming, T. R. 'Intermediate-sized trials for the evaluation of HIV vaccine candidates (abstract)., in *Conference on advances in AIDS vaccine development: 8th Annual Cooperative Vaccine Development Groups for AIDS Meeting Program book*, National Institutes of Health, Bethsada, 1996, p. 108.
3. Halloran, M. E., Haber, M. J., Longini, I. M. and Struchiner, C. J. 'Direct and indirect effects in vaccine field efficacy and effectiveness', *American Journal of Epidemiology*, **133**, 323–331 (1991).
4. Halloran, M. E., Struchiner, C. J. and Spielman, A. 'Modeling malaria vaccines II: Population effects of stage-specific malaria vaccines dependent on natural boosting', *Mathematical Biosciences*, **94**, 115–149 (1989).
5. Halloran, M. E., Cochi, S., Lieu, T., Wharton, M. and Fehrs, L. 'Theoretical epidemiologic and morbidity effects of routine immunization of preschool children with live-virus varicella vaccine in the U.S.', *American Journal of Epidemiology*, **140**, 81–104 (1994).
6. Halloran, M. E., Longini, I. M., Struchiner, C. J., Haber, M. J. and Brunet, R. C. 'Exposure efficacy and change in contact rates in evaluating prophylatic HIV vaccines in the field', *Statistics in Medicine*, **13**, 357–377 (1994).
7. Halloran, M. E. and Struchiner, C. J. 'Causal inference for infectious diseases', *Epidemiology*, **61**, 142–151 (1995).
8. Halloran, M. E. 'Evaluating HIV vaccines: Discussion', *Statistics in Medicine*, **15**, 2405–2412 (1996).
9. DeGruttola, V., Seage, G. R., Mayer, K. H. and Horsburgh, C. R. 'Infectiousness of HIV between male homosexual partners', *Journal of Clinical Epidemiology*, **42**, 849–856 (1989).
10. Wiley, J. A., Herschkorn, S. J. and Padian, N. S. 'Heterogeneity in the probability of HIV transmission per sexual contact: the case of male-to-female transmission in penile-vaginal intercourse', *Statistics in Medicine*, **8**, 93–102 (1989).

11. Longini, I. M., Clark, W. S., Haber, M. J. and Horsburgh, R. C. 'The stages of HIV infection: Waiting times and infection transmission probabilities', *Lecture Notes in Biomathematics*, **83**, 111–137 (1989).
12. Kim, M. Y. and Lagakos, S. W. 'Estimating the infectivity of HIV from partner studies', *Annals of Epidemiology*, **1**, 117–128 (1990).
13. Jewell, N. P. and Shiboski, S. C. 'Statistical analysis of HIV infectivity based on partner study data', *Biometrics*, **46**, 1133–1150 (1990).
14. Shiboski, S. C. and Jewell, N. P. 'Statistical analysis of the time-dependence on HIV infectivity based on partner study data', *Journal of the American Statistical Association*, **87**, 360–372 (1992).
15. Koopman, J. S. and Little, R. J. 'Assessing HIV vaccine effects', *American Journal of Epidermiology*, **142**, 1113–1120 (1995).
16. Longini, I. M. and Koopman, J. S. 'Household and community transmission parameters from final distributions of infections in households', *Biometrics*, **38**, 115–126 (1982).
17. Rida, W. N. 'Assessing the effect on HIV vaccination on secondary transmission', *Statistics in Medicine*, **15**, 2393–2404 (1996).
18. Longini, I. M., Datta, S. and Halloran, M. E. 'Measuring vaccine efficacy for both susceptibility to infection and reduction in infectiousness for prophylactic HIV-1 vaccines', *Journal of Acquired Immune Deficiency Syndrome and Human Retrovirus*, **13**, 440–447 (1996).
19. Halloran, M. E. and Struchiner, C. J. 'Evaluating vaccine effects: changes in susceptibility, infectiousness, contacts, direct and indirect effects', in *Proceeding of the III Brazilian/II Ibero American/Latin American Congress on Epidemiology*, April, 1995.
20. Rhodes, P. H., Halloran, M. E. and Longini, I. M. 'Counting process models for differentiating exposure to infection and susceptibility', *Journal of the Royal Statistical Society, Series B*, **58**, 751–762 (1996).
21. Greenwood, M. and Yule, U. G. 'The statistics of anti-typhoid and anti-cholera inoculations, and the interpretation of such statistics in general', *Proceedings of the Royal Society of Medicine*, **8** (part 2), 113–194 (1915).
22. Longini, I. M. and Halloran, M. E. 'A frailty mixture model for estimating vaccine efficacy', *Applied Statistics*, **45**, 165–173 (1996).
23. Halloran, M. E., Longini, I. M. and Struchiner, C. J. 'Estimability and interpretation of vaccine efficacy using frailty mixing models', *American Journal of Epidemiology*, **144**, 83–97 (1996).
24. Magder, L. and Brookmeyer, R. 'Analysis of infectious disease data from partners studies with unknown source of infection', *Biometrics*, **49**, 1110–1116 (1993).
25. Smith, P. G., Rodrigues, L. C. and Fine, P. E. M. 'Assessment of the protective efficacy of vaccine against common diseases using case-control and cohort studies', *International Journal of Epidemiology*, **13**, 87–93 (1984).
26. Halloran, M. E., Haber, M. J. and Longini, I. M. 'Interpretation and estimation of vaccine efficacy under heterogeneity', *American Journal of Epidemiology*, **136**, 328–343 (1992).
27. Brunet, R. C., Struchiner, C. J. and Halloran, M. E. 'On the distribution of vaccine protection under heterogeneous response', *Mathematical Biosciences*, **116**, 111–125 (1993).
28. Cohen, J. 'News and comment: Thailand weighs AIDS vaccine tests', *Science*, **270**, 904–970 (1995).
29. Addy, C. L., Longini, I. M. and Haber, M. J. 'A generalized stochastic model for the analysis of infectious disease final size data', *Biometrics*, **47**, 961–974 (1991).

.