Proceedings of the 2000 Winter Simulation Conference
J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, eds.

MONTE CARLO SIMULATION EXPERIMENTS FOR ANALYSIS OF HIV VACCINE EFFECTS AND VACCINE TRIAL DESIGN

Daniel C. Barth-Jones

Center for
Healthcare Effectiveness Research
Wayne State University,
School of Medicine
121 Shiffman Medical Library,
4235 Brush Street
Detroit, MI 48201, U.S.A

Andrew L. Adams

NextHop Technologies
4251 Plymouth Road
Building 1, Suite 2000
Ann Arbor, MI 48105, U.S.A.

James S. Koopman

Department of Epidemiology
School of Public Health,
University of Michigan
109 Observatory Street
Ann Arbor, MI 48109, U.S.A.

ABSTRACT

The field of infectious disease epidemiology has increasingly adopted stochastic simulation technologies to simulate complex infectious disease transmission systems. Such simulations have both increased the scientific understanding of infectious disease transmission dynamics and served as important tools for evaluating epidemiologic study designs and statistical methods. This paper reports on a discrete-event simulation to analyze the recently developed Retrospective Partner Trials (RPT) HIV vaccine trial design. A specially designed simulation system, HIVSIM, was used to simulate data resulting from the RPT design vaccine trials. HIVSIM explicitly models complex HIV transmission dynamics (e.g., sexual partner mixing patterns and concurrent sexual partnerships) and vaccine trial design characteristics. Monte Carlo simulation analyses conducted with HIVSIM indicate that the RPT design is able to produce vaccine effect estimates with acceptably small bias, high precision and excellent statistical power under plausible HIV vaccine trial conditions. Additionally, the explicit simulation of HIV transmission dynamics permits investigations into the common, but unwarranted, statistical independence assumptions routinely used in the estimation of vaccine effects.

1 INTRODUCTION

1.1 Simulation Background

Within the field of infectious disease epidemiology, computer simulation techniques have increasingly been employed to model the spread of infectious diseases within human populations (Ackerman 1984, Mollison 1995, Isham and Medley 1996). These dynamic epidemic simulations are typically used to advance scientific understanding of how the interactions of individuals create a transmission system that allows the infectious agent to spread within a human population. Such simulations are also increasingly employed for the purposes of evaluating epidemiologic study designs and statistical methods (Boily and Anderson 1996, Barth-Jones 1999)

This paper reports on the use of one such simulation to evaluate an HIV vaccine trial design. The evaluation of the vaccine trial design was accomplished by performing a Monte Carlo simulation experiment. The data for the experiment was generated by HIVSIM, a computer program that simulates HIV vaccine trials conducted in a homosexual mixing population. HIVSIM explicitly models characteristics of the HIV transmission system and HIV vaccine trial design that could not be easily investigated using mathematical models and differential equation-based compartmental models (Anderson and May 1991, Jacquez 1996, Adams et al. 1998). The experiment reported here is distinguished from more typical biostatistical Monte Carlo experiments by the fact that explicit simulation of causal processes produced the experimental outcome events. In contrast, typical Monte Carlo experiments commonly draw outcome events from known standard probability distributions in order to determine the sampling distributions of the parameters being estimated (Mooney 1997). Although Monte Carlo experiments in which transmission dynamics are explicitly simulated are not new (Fox et al. 1971), such experiments have recently seen increased use. The explicit simulation of transmission dynamics has recently allowed researchers to address statistical estimation and study design issues related to: 1) sexually transmitted diseases (STDs) as risk factors for HIV infection (Boily and Anderson 1996); 2) the effects of social networks on the spread of STDs (Ghani et al. 1998); and 3) HIV vaccine effect estimation (Longini et al. 1998, Barth-Jones et al. 1998, Desai et al. 1999).

1.2 Epidemiologic Background

In the more than 15 years since the discovery of HIV, the HIV/AIDS pandemic has continued to expand relentlessly and has now become the fourth leading cause of death in the world and the leading cause of death on the continent of Africa (Balter 1999). The United Nations AIDS program, UNAIDS, has estimated that at least as many as 6 million new HIV infections are now occurring each year (Balter 1999). Such statistics highlight the vital importance that an effective prophylactic HIV vaccine could have in achieving control over this worldwide epidemic. Toward this end, Phase III clinical trials to assess the protective efficacy of a candidate HIV vaccine based on recombinant envelope glycoprotein (gp) 120 began in 1998 in the United States and in 1999 in Thailand (Francis et al. 1998).

However, HIV vaccines may not fully protect against HIV infection in all individuals. If this proves to be the case, then it will be important to recognize that even vaccines that fail to prevent HIV infection in all individuals could have beneficial effects. On an individual level, the vaccine could delay or prevent the onset of AIDS; while on the population level, the vaccine might reduce the infectiousness to others for individuals who are infected despite vaccination (Koopman and Little 1995).

Such issues have motivated several recent publications in which vaccine effects that reduce the infectiousness of those vaccinated are addressed (Longini et al. 1996, Rida 1996, Datta et al. 1998, Longini et al. 1999, Barth-Jones 1999). Vaccine effects resulting in the reduction of the infectiousness of vaccinated individuals are termed Vaccine Effects on Infectiousness (VE_I). Vaccine effects that reduce the susceptibility of an individual to infection are termed Vaccine Effects on Susceptibility (VE_S). VE_S and VE_I can be considered as proportional reductions in the susceptibility or infectiousness, respectively, of vaccinated individuals due to the effects of the vaccine. Typically, VE_S is formulated as 1-RR, where RR is the relative risk of becoming infected in vaccinated individuals as compared to unvaccinated individuals. For the VE_I, the vaccine effect measure would take the form of a ratio of two conditional secondary attack rates (SARs) (i.e., SAR_V/SAR_P), where SAR_V would indicate the fraction of individuals who were infected by exposure to an infected vaccinated individual, and SAR_P would indicate the fraction of individuals who were infected by exposure to an infected unvaccinated individual.

Vaccine Effects on Infectiousness have important public health implications for the ability of a vaccine to disrupt epidemic transmission dynamics, particularly if VE_S effects are not substantial. Vaccines which only modestly reduce susceptibility to HIV infection, but which also reduce the infectiousness of vaccinated individuals, could potentially halt epidemic HIV transmission primarily by preventing infected individuals from transmitting HIV to others (Adams et al. 1998).

The combined effects of VE_S and VE_I on the epidemic potential of a disease can be summarized under a set of hypothetical conditions by another vaccine efficacy measure called VE_R, the Vaccine Effect on the basic Reproduction number (Adams et al. 1998). The Basic Reproduction Number (R_0) can be understood as the average number of secondary infections produced when one infected individual is introduced into a population where everyone is susceptible (Anderson and May 1991). When R_0 is greater than one, epidemic disease transmission results; when R_0 is less than one, disease transmission cannot be sustained and eventually the disease will become eradicated. VE_R can be defined as:

$$VE_R = 1-((1-VE_S)\,(1-VE_I)), \qquad (1)$$

and can be interpreted as the proportional reduction in the basic reproduction number caused by vaccination under a set of hypothetical conditions which would hold in a simple mathematical model of epidemic disease transmission (Barth-Jones 1999). Although these hypothetical conditions are unlikely to exactly hold for HIV (or other infectious diseases), this definition of VE_R provides a convenient framework for expressing the potential combined effects of VE_S and VE_I on the transmission of an infectious agent within a population.

The Retrospective Partner Trials (RPT) HIV vaccine trial design was developed to allow the measurement of VE_S, VE_I, and VE_R. Standard vaccine trial designs only measure VE_S, and, thus, run the risk of rejecting these potentially effective HIV vaccines. The RPT vaccine trial design obtains information on VE_I by employing contact tracing (also known as partner notification) to follow the sex partners of those vaccine trial participants who become infected during the vaccine trial (Adams et al. 1998, Barth-Jones 1999). Recent work by Barth-Jones and colleagues describes the RPT vaccine trial design; presents methods for estimating VE_S, VE_I and VE_R within the design; and explores the implications of a probability model for the statistical power of the RPT design (Adams et al. 1998, Barth-Jones 1999). Like most simple methods for estimating vaccine effects, all of these methods are based on the normal approximation to the binomial distribution (O'Neill 1988, Blackwelder 1993).

An aspect of the RPT design that requires investigation involves the manner in which the RPT design calculates the SARs used to estimate VE_I and, thus, VE_R. Perfect calculation of the SARs would involve exact knowledge of the time of HIV infection for each of the vaccine study participants. The RPT trial design, however, must approximate these times of infection using information on the timing of the sexual partnerships of infected study participants and phylogenetic information from HIV collected from the infected study participants and their sex partners. Using simulation data, it is possible to compare the effects of the imperfect RPT SAR calculation methods with the

true VE_I and VE_R results, which can be collected in the computer simulation but which would be unascertainable in an actual vaccine trial.

The Monte Carlo experiment presented here has three main goals. The first is to investigate the bias, precision and statistical power of the vaccine effect estimators for VE_S, VE_I and VE_R. The second is to compare the VE_I and VE_R estimators produced by the RPT design with the VE_I and VE_R estimators produced using perfect information from the simulation. The final goal is to demonstrate that infectious disease transmission dynamics violate the independence assumptions inherent in the use of the binomial distribution for modeling infection risks.

Section 2 of this report briefly describes the study design for an RPT HIV vaccine trial. Section 3 provides a summary of the HIVSIM computer simulation used to generate the simulated vaccine trial data. The design of the Monte Carlo simulation experiment is presented in Section 4. Section 5 provides the results from the Monte Carlo experiment. Finally, Section 6 discusses the implications of these results for the RPT trial design and the important role simulation methods have had in the development of this HIV vaccine trial design.

2 RETROSPECTIVE PARTNER TRIALS STUDY DESIGN

The RPT design employs retrospective contact tracing to obtain information regarding the sex partners of those trial subjects who became infected during the vaccine trial. This information is used in the calculation of the VE_I estimates. Two types of information obtainable during the contact tracing process are used to calculate the SARs used to estimate VE_I: 1) phylogenetic analysis of blood samples from the trial subjects and their sex partners (Leitner et al. 1996), and 2) the timing of the sexual partnerships. See Adams et al. (1998) and Barth-Jones (1999) for a more thorough discussion.

The RPT trial would be conducted as a double-blind, placebo controlled randomized vaccine trial. A total sample size of N_{tot} HIV negative vaccine trial participants would be enrolled into the vaccine trial. A fraction f_v of the vaccine trial participants would be randomized to the vaccine arm of the trial (N_v) and the fraction 1- f_v would be randomized to the placebo arm of the trial (N_p). After receiving either vaccine or a placebo, participants would then be followed with periodic testing for HIV infection. Those trial participants who are found to be infected would be considered index cases. Vaccinated index cases (I_v) and placebo-treated index cases (I_p) would be interviewed to provide information for the contact-tracing process. Sex partners identified via the contact-tracing process would be notified of potential exposure to an HIV infected partner and offered HIV testing. Those partners to whom transmission could potentially have occurred would be determined on the basis of phylogenetic analysis and historical timing data. Such partnerships constitute "trial" partnerships. The total number of these trial partnerships for the vaccinated (T_v) and placebo-treated (T_p) index cases serve as the denominators in the SARs for the calculation of VE_I. All partnerships with related HIV which began subsequent to the end of the first partnership found to have HIV related to that of the index case would be considered transmission partnerships. The total number of transmission cases for the vaccinated (X_v) and placebo-treated (X_p) index cases would serve as the numerators in the SARs for VE_I.

3 COMPUTER SIMULATION METHODS

As previously mentioned, HIVSIM, a discrete event computer simulation (DES), was employed for the evaluation of the RPT study design because it can generate simulated vaccine trial data which are produced by a complex, dynamic HIV transmission model. HIVSIM was constructed as a Monte Carlo DES in a continuous time framework with randomly determined times to future events. This simulation environment is capable of generating many real world complexities of HIV transmission dynamics and HIV vaccine trials. Real world sexual contact patterns can include partnerships of variable lengths that might be made preferentially between different types of individuals depending on the partnership status of those individuals. Additional characteristics of real world contact patterns and HIV transmission include: 1) concurrent as well as monogamous partnerships; 2) partnerships with differing rates of sexual activity; 3) individuals with different partnership-seeking propensities; and 4) HIV transmission probabilities which are dependent on sex-act type, sex-act role and stage of HIV infection. All of these behaviors/ conditions can be manipulated by modifying the HIVSIM simulation parameters. These simulation parameters determine the expectations for the stochastic variation that occurs in each individual within each simulation run. Only a brief overview of the HIVSIM vaccine trial simulation process is provided here. The HIVSIM computer simulation model design, function and implementation details is described in Adams et al. 1998.

Prior to beginning a vaccine trial experiment, HIVSIM runs through a partnership stabilization period in order to allow a stable partnership contact pattern to develop. Once this equilibration has occurred, all simulation repetitions for an experiment are started from the same initial partnership formation state. At the start of each simulation repetition, a number of individuals are randomly selected from the population and infected with HIV. Immediately thereafter, a simulated vaccine trial is initiated by: 1) the random selection of vaccine trial participants within the population, and 2) their randomization to the vaccine or placebo arms of the vaccine trial. The sexual mixing and the HIV transmission processes continue throughout the simulated

vaccine trial period. HIV transmission probabilities within the simulation are specified according to type of sex-acts, sexual roles, and the stage of HIV infection for the infected partner. In vaccinated individuals, these transmission probabilities are modified according to the VE_S and VE_I parameters specified for the simulation. For example, if an uninfected partner has been vaccinated, the transmission probability from the infected participant to the uninfected partner is multiplied by a proportional factor of 1-VE_S. Likewise, if the infected partner has become infected despite vaccination, the transmission probability from the infected participant to an uninfected partner is multiplied by a proportional factor of 1-VE_I. Finally, if both partners have been vaccinated, then the transmission probability between the infected and the uninfected partner is multiplied by a proportional factor equal to the product of 1-VE_I and 1-VE_S.

Complete information is recorded on all sexual partnerships and HIV transmission events during the simulated vaccine trial. At the end of the specified vaccine trial period, the vaccine efficacy measurements (VE_S, VE_I and VE_R) from the trial are calculated. This allows the calculation of two separate types of estimates for VE_I from the simulated population. "RPT-based" vaccine efficacy measurements utilize the same information that would be available in an actual RPT vaccine trial. These measurements rely on classification rules to establish which sexual partnerships of an index case can be considered as part of the numerator or the denominator in the SARs used to calculate VE_I (Adams et al. 1998). "Simulation-based" estimates utilize perfect information that is available only in the simulated environment. In the computer-simulated environment, the exact time of HIV infection is known for every individual in the vaccine trial population. This allows simulation-based VE_I estimates to perfectly determine whether each sexual partnership of an index case in the vaccine trial should constitute a "trial" partnership that was at risk for HIV transmission from the index case. RPT-based VE_I estimates, however, must rely on the same information which would be obtainable in an actual RPT vaccine trial and, therefore, will be subject to some imperfection in ascertaining which sexual partnerships of an index case actually constitute trial partnerships. Therefore, it is of interest to compare the precision and bias of RPT-based and simulation-based VE_I estimates produced by the HIVSIM simulation of the RPT vaccine trial. Because VE_R estimates result from combined VE_S and VE_I estimates, VE_R estimates can also be designated as RPT-based or simulation-based estimates depending on the type of VE_I estimates from which they were derived.

4 MONTE CARLO EXPERIMENT DESIGN AND METHODS

A Monte Carlo simulation experiment was performed using HIVSIM to simulate repeated RPT vaccine trials with identical conditions. Each simulation repetition provides one estimate for each of the observed statistics drawn from their respective sampling distributions.

Three different vaccine effect settings were selected for this experiment. The VE_S and VE_I parameters and the number of simulated vaccine trials conducted for each setting were as follows: 1) VE_S=0.0, VE_I=0.0, R=1,000 simulated vaccine trials; 2) VE_S=0.25, VE_I=0.75, R=500; 3) VE_S=0.5, VE_I=0.9, R=500. All other model parameters for HIVSIM were identical across the three vaccine effect settings and across the vaccine trial repetitions. One thousand repetitions were conducted for the VE_S=0.0, VE_I=0.0 setting in order to assure well-defined null distribution tails for empirical statistical power calculations.

All of the simulations were conducted with identical starting conditions for the partnership network in the sexual mixing population as described in the Computer Simulation Methods section. Each simulated multi-site vaccine trial took place within 4 closed homosexual mixing populations of 4,400 individuals. The RPT vaccine trials were each conducted over a simulated period of two years. Individuals could leave the population only through death by AIDS. The simulation parameters were set so that 25 percent of the sexual mixing population would participate in concurrent sexual partnerships with as many as two simultaneous sexual partnerships. Prior to the initiation of the vaccine trial, 400 individuals within each of the four study site populations were randomly selected for HIV infection. Of the remaining 4,000 uninfected members in each of the four simulated populations, a total of 660 individuals were randomly selected for participation in each RPT vaccine trial. The total sample size for the multi-site RPT vaccine trial was thus N=2,640, with N=1,320 in both the vaccinated and the placebo treatment groups. This sample size is roughly comparable to the actual sample size of the Phase III HIV vaccine trials which are currently being conducted in Thailand with a sample size of N=2500 (Francis 1998, Balter 1998). The sample size within each study site population was selected so that only a small percentage (7.5%) of the mixing population would be vaccinated. This proportion was small enough that the epidemic transmission in the mixing population was not significantly altered due to herd immunity or indirect vaccine effects. Barth-Jones (1999) provides a complete list of the HIVSIM simulation parameters used in this experiment.

Once all of the simulation repetitions are completed, the results of the simulation experiment can be summarized. The following notation will be used to summarize the results of the simulation experiment. Let Λ be a general symbol for any simulation parameter that was set to a specific value for a series of simulation repetitions. Let the total number of simulation repetitions be denoted as R, and the simulation repetitions be indexed by j. Then, for any general statistic, $\hat{\Lambda}_j$ will denote the realization of a single estimate of the simulation parameter taken from the

*j*th simulation. The sample mean of the *R*, $\hat{\Lambda}_j$ Monte Carlo estimates for the simulation parameter Λ will be denoted as $\overline{\Lambda}$. For a given set of *R* simulation repetitions, then, the Monte Carlo estimate of the bias of the statistic relative to the simulation parameter is given by:

$$\hat{BIAS}\,[\hat{\Lambda}] = \overline{\Lambda} - \Lambda\,. \qquad (2)$$

The bias of a statistic indicates, on average, how much the estimator will over- or underestimate the actual parameter value (Mooney 1997). The Monte Carlo estimate of the Standard Error of the statistic is given by:

$$\hat{SE}\,[\hat{\Lambda}] = \sum_{j=1 to R} (\hat{\Lambda}_j - \overline{\Lambda})^2 / (R-1)\,. \qquad (3)$$

The standard error indicates the precision with which a statistic is estimated (Mooney 1997). The Monte Carlo estimate of the Root Mean Square Error of the statistic is given by:

$$\hat{RMSE}\,[\hat{\Lambda}] = \sum_{j=1 to R} (\hat{\Lambda}_j - \Lambda)^2 / R\,. \qquad (4)$$

The Root Mean Square Error of a statistic is a measure of its accuracy that takes into account both the bias and the standard error. The ratio of the RMSE over the standard error indicates the relative increase in the overall error that is caused by the bias. Typically, unless it is desired to have highly accurate confidence intervals for a statistic, if the bias contributes only a few percentage points of increase to the overall error relative to the standard error, then it is ignorable (Efron 1993).

The precision of the Monte Carlo estimates made with Equations (2-4) will increase with an increasing number of simulation repetitions. However, because HIVSIM simulation runs are very computationally intensive, the precision of the Monte Carlo estimates was further increased by the use of bootstrap methods (Efron 1993). The bootstrap estimates were made by drawing samples of size *R* with replacement from the original *R* Monte Carlo simulation experiments. A total of 2,000 bootstrap samples were drawn to produce the bootstrapped estimates of the Monte Carlo results. The bootstrapped estimates of the bias, MSE and SE are the average values for Equations (2-4), respectively, applied to the 2,000 bootstrap samples of the original Monte Carlo results. Although the bootstrapped estimates are reported here, these estimates differed from the original Monte Carlo estimates only in the fourth and, rarely, the third digits, indicating that the precision of the original Monte Carlo estimates was quite high. The Monte Carlo estimates for the equitailed 1-α percent confidence intervals were obtained by determining the bootstrap estimates for the $\alpha/2$ and 1-$\alpha/2$ percentiles of the distribution of the *R* Monte Carlo estimates, Λ_j (Mooney 1997).

The empirical statistical power was estimated as follows. First, the bootstrap estimate for the 95th percentile of the null distribution (i.e., simulation parameters: VE_S=0.0 and VE_I=0.0) was determined. Then the empirical statistical power for a one-tailed, α=0.05 test was calculated as the proportion of the R simulation repetitions in the effect distributions (i.e., simulation parameters $VE_S > 0.0$ and $VE_I > 0.0$), where the Monte Carlo estimates exceeded the null distribution's bootstrap 95th percentile point. (Mooney 1997). Statistical power was also determined for the large sample Wald tests and the permutation tests for VE_S, VE_I and VE_R (Blackwelder 1993, Barth-Jones 1999, Good 1994). The statistical power estimates constituted the proportion of the *R* simulation repetitions for the effect distributions which had Wald (or permutation) tests which were statistically significant at the one-tailed, α=0.05 level.

5 RESULTS

5.1 Sampling Distribution Characteristics

Table 1 provides a summary of simulation experiment results for the sampling distribution characteristics of the vaccine efficacy parameters VE_S, VE_I and VE_R. The sampling bias and precision for these estimators are summarized by the bootstrap estimates for the Bias (Equation 2), the Standard Error (SE) (Equation 3) and the Root Mean Square Error (RMSE) (Equation 4). The values presented in Table 1 are bootstrap estimates of the Monte Carlo sampling distribution estimates as described in the Monte Carlo Experimental Design and Methods section. Each of the VE measurements: VE_S, $VE_{I\ RPT}$, $VE_{I\ SIM}$, $VE_{R\ RPT}$ and $VE_{R\ SIM}$, are arranged in blocks of rows for the purpose of comparison. Two types of comparisons are of primary interest for the results presented in this table. First, the changes in the bias and precision for each VE statistic can be observed as the VE_S and VE_I parameters were altered in the simulation experiment. Second, for VE_I, the changes in bias and precision resulting from the imperfect estimation of the time of infection for the index cases can be assessed by comparing the $VE_{I\ RPT}$ and $VE_{I\ SIM}$ results. Likewise, comparing $VE_{R\ RPT}$ with $VE_{R\ SIM}$ results can assess the impact of this imperfect VE_I estimation on VE_R.

All of these statistics demonstrate a negative bias, indicating that on average they will underestimate the actual vaccine efficacy statistics. For VE_S, the bias increases from -0.02, when VE_S = 0.0, to approximately –0.06, when VE_S = 0.5. For the other four statistics, $VE_{I\ RPT}$, $VE_{I\ SIM}$, $VE_{R\ RPT}$ and $VE_{R\ SIM}$, the bias is greatest when the vaccine efficacy parameters are 0.0, but decreases considerably with the increase in the VE statistics. It is also interesting to note that the bias is slightly larger for the simulation measures $VE_{I\ SIM}$ and $VE_{R\ SIM}$ than it is for the

Table 1: Monte-Carlo Simulation Experiment Results for Sampling Distribution Characteristics

(1) Simulation Parameter Settings	(2) Simulation Repetitions (R =)	(3) VE Parameter Under Consideration	(4) Mean	(5) Bias	(6) SE	(7) RMSE	(8) Ratio RMSE/ SE	(9) Low 90% CI (5% below value)	(10) High 90% CI (5% above value)
		VE_S							
VE_S=0, VE_I=0	1,000	0	-0.002	-0.002	0.149	0.149	1.000	-0.267	0.225
VE_S=0.25, VE_I=0.75	500	0.25	0.217	-0.033	0.136	0.140	1.029	-0.019	0.416
VE_S=0.5, VE_I=0.9	500	0.5	0.443	-0.057	0.104	0.120	1.143	0.260	0.600
		$VE_{I\ RPT}$							
VE_S=0, VE_I=0	1,000	0	-0.057	-0.057	0.304	0.310	1.017	-0.612	0.354
VE_S=0.25, VE_I=0.75	500	0.75	0.7163	-0.034	0.140	0.144	1.029	0.464	0.896
VE_S=0.5, VE_I=0.9	500	0.9	0.8847	-0.015	0.101	0.102	1.012	0.700	1.000
		$VE_{I\ SIM}$							
VE_S=0, VE_I=0	1,000	0	-0.045	-0.045	0.258	0.262	1.015	-0.512	0.317
VE_S=0.25, VE_I=0.75	500	0.75	0.6985	-0.052	0.124	0.135	1.083	0.475	0.866
VE_S=0.5, VE_I=0.9	500	0.9	0.8726	-0.027	0.091	0.095	1.044	0.706	1.000
		$VE_{R\ RPT}$							
VE_S=0, VE_I=0	1,000	0	-0.060	-0.060	0.351	0.356	1.014	-0.710	0.401
VE_S=0.25, VE_I=0.75	500	0.8125	0.778	-0.034	0.114	0.119	1.044	0.565	0.924
VE_S=0.5, VE_I=0.9	500	0.95	0.936	-0.014	0.057	0.059	1.030	0.831	1.000
		$VE_{R\ SIM}$							
VE_S=0, VE_I=0	1,000	0	-0.048	-0.048	0.309	0.312	1.012	-0.617	0.370
VE_S=0.25, VE_I=0.75	500	0.8125	0.764	-0.048	0.104	0.114	1.103	0.570	0.902
VE_S=0.5, VE_I=0.9	500	0.95	0.929	-0.021	0.052	0.056	1.078	0.834	1.000

equivalent RPT-based measures. However, when the VE_I and VE_R parameters are high, the observed biases for both types of estimators become smaller and the differences in the degree of bias between the RPT-based and simulation-based VE statistics decreases. For example, for the parameters VE_S=0.5 and VE_I=0.9, the respective biases for $VE_{I\ RPT}$ and $VE_{I\ SIM}$ are only -0.015 and -0.027.

The standard errors of the vaccine efficacy statistics are acceptably small for all of the estimators at the non-null experimental parameter settings. The standard errors for all of the VE statistics decrease with the increasing vaccine effects. As would be expected, the standard errors for the RPT estimates are consistently higher than the simulation estimates. These same patterns are also seen in the 90 percent confidence intervals presented in Columns 9 and 10.

The Root Mean Square Errors demonstrate a similar trend for all of the VE statistics as was observed for the standard errors. The RMSE indicates how both the bias and the standard error affect the accuracy of the VE statistics. The RMSE for all of the VE statistics decreases with the increasing vaccine effects. As was found for the standard errors, the RMSEs for the RPT-based estimates for VE_I and VE_R are consistently slightly higher than the simulation-based estimates. Because the exact time of HIV infection is imperfectly approximated by the RPT-based estimates, these estimates have slightly decreased precision relative to the simulation-based estimates. This conclusion is supported by the modest values found in Column 8 for the ratios of the RMSEs over the standard errors. The largest ratio, 1.14, occurs for the VE_S statistic when the VE_S parameter is 0.5. Because the RMSE is only about 14 percent greater than the standard error alone, it can be concluded that bias has a small, but not ignorable, influence on the overall accuracy of these statistics. The observed bias in these estimates can be mostly accounted for by issues related to well-known problem of bias of in ratio estimators. Chick et al. (In Press) investigate this issue for vaccine effect estimators and present bias corrections for VE_S, VE_I and VE_R. These simulation data were also reanalyzed using a more appropriate estimator for vaccines that only provide partial protection against infection or transmission (Haber et al. 1991), and the bias correction of Jewell (1986) or its VE_I and VE_R analogs (Chick et al. In Press). With these improved estimators, the bias was found to be reduced to almost one-half of the levels reported here. Additionally, the precision of these estimators was also improved in most cases (data not shown).

5.2 Statistical Power Comparisons

Table 2 provides the results comparing the empirical statistical power with the power of the Wald and permutation hypothesis testing methods. All of the statistical power calculations are reported for a one-tailed test with α=0.05.

The most important result from this investigation of the statistical power is the fact that the simulations demonstrated substantial statistical power for the RPT design to detect VE_I and VE_R effects under large sample conditions. For example, both the Wald and permutation tests demonstrated statistical power above 90 percent for $VE_{I\ RPT}$, when VE_S=0.25 and VE_I=0.75 (i.e., VE_R = 0.8125). The statistical power for $VE_{R\ RPT}$ under the same conditions was also over 90 percent for both the Wald and the permutation tests. Note that a vaccine with a VE_R greater than 80 percent, would be theoretically capable of reducing an R_0 of 5.0 – the upper estimate for the R_0 of the AIDS epidemic in San Francisco (Blower and McLean 1994) - to below one, thus halting epidemic HIV transmission. By comparison, a standard vaccine trial design, which is only capable of detecting VE_S effects, would have had only a 49.2 percent probability of detecting the VE_S effect using the Wald test. Therefore, a standard vaccine trial would have had a 50 percent chance of rejecting an HIV vaccine with the theoretical potential to halt the San Francisco AIDS epidemic.

It should also be noted that in some cases the power for the Wald and Permutation tests under the null distribution exceeds the nominal alpha level for the tests. This means that the Wald and Permutation tests will make more type I errors than were specified as acceptable. Theoretically, the permutation test results should be conservative, with the observed alpha levels always falling at or below the nominal alpha level, if the independence and randomization assumptions of the test have been met (Good 1994). The failure of the permutation tests to yield conservative or exact alpha errors is thought by the authors to be caused by the lack of independence of outcomes for the participants in the vaccine trials. While the permutation test dispenses with the assumption of the binomial distribution, the violation of the independence assumption still appears to be an important issue for the use of the permutation test in this context. The significance of the violation of the independence assumptions is discussed in more detail in the following section the on correlation of infection risks.

5.3 Correlation of Infection Risks and Independence of Observations

Figure 1 provides a scatterplot of the infection risks in the vaccinated and placebo treatment groups for each of the 4000 simulated study sites that were used to form the 1,000 multi-site vaccine trials where VE_S=0 and VE_I=0.

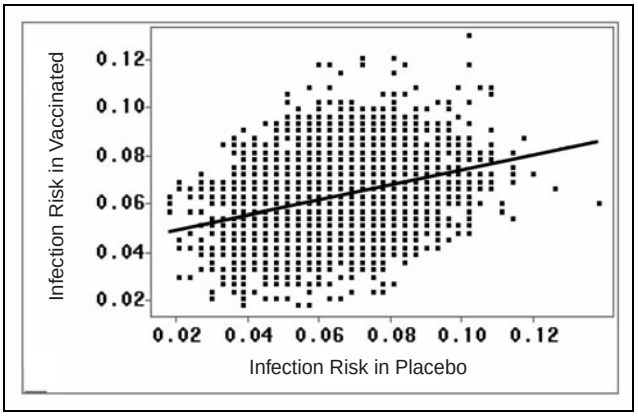

Figure 1: Correlation in Infection Risks

Table 2: Summary of Statistical Power Comparisons

(1) Simulation Parameter Settings	(2) Empirical Statistical Power (%)	(3) Statistical Power for Wald Hypothesis Tests (%)	(4) Statistical Power for Permutation Hypothesis Tests (%)	(5) Empirical Statistical Power (%)	(6) Statistical Power for Wald Hypothesis Tests (%)	(7) Statistical Power for Permutation Hypothesis Tests (%)
	Statistical Power for VE_S					
VE_S=0, VE_I=0, (VE_R=0)	5.0	4.9	8.5			
VE_S=0.25, VE_I=0.75, (VE_R=0.8125)	42.2	49.2	43.6			
VE_S=0.5, VE_I=0.9, (VE_R=0.95)	91.2	93.6	89.5			
	Statistical Power for $VE_{I\ RPT}$			Statistical Power for $VE_{I\ SIM}$		
VE_S=0, VE_I=0, (VE_R=0)	5.0	8.6	7.5	5.0	8.2	8.6
VE_S=0.25, VE_I=0.75, (VE_R=0.8125)	93.6	94.0	91.0	95.6	96.6	93.0
VE_S=0.5, VE_I=0.9, (VE_R=0.95)	98.4	79.7	96.8	99.2	92.0	98.4
	Statistical Power for $VE_{R\ RPT}$			Statistical Power for $VE_{R\ SIM}$		
VE_S=0, VE_I=0, (VE_R=0)	5.0	3.1	7.4	5.0	2.6	8.8
VE_S=0.25, VE_I=0.75, (VE_R=0.8125)	95.6	93.2	94.4	97.2	96.6	95.8
VE_S=0.5, VE_I=0.9, (VE_R=0.95)	100.0	80.9	99.8	100.0	92.4	100.0

As can be seen from the regression line on the graph, a modest correlation exists between the infection risks in the vaccinated treatment group and the placebo treatment group. The correlation coefficient for these risks was 0.32. Moderate, but highly significant, correlations in the infection risks were also were seen for the other parameter settings used in this experiment. Epidemic transmission was the cause of the correlation between these risks. Although each study site simulation was conducted with identical starting conditions, stochastic variations in the evolving HIV epidemics caused the infection risks to vary in each simulation. This is the case because the infection risk for each individual in an epidemic population is dependent on the probability that they will make contact with an infected individual. As each epidemic moves through the simulated populations, the prevalence of HIV in the populations increases, as does the probability that the population members will subsequently make contact with an infectious sexual partner. This causes the infection risks of the population members to be interdependent, thus violating the independence assumption inherent in the use of the Binomial probability distribution (Collet 1994).

Within a single real-world vaccine trial or observational study of infectious disease risks, the correlation of binary responses cannot be easily demonstrated or observed. However, within a simulation environment, where epidemics can be observed many times under identical operating conditions, the correlation of the binary responses is readily observed, as it is in Figure 1.

6 DISCUSSION

This investigation used a computer simulation model of HIV transmission within a homosexual mixing population to generate simulated data from an RPT HIV vaccine trial. The computer simulation data was used to: 1) examine the sampling distribution characteristics and statistical power for the VE_S, VE_I and VE_R estimates; 2) demonstrate that the imperfect RPT methods for measuring VE_I effects do not significantly reduce the accuracy of the VE_I and VE_R estimates made by the RPT trials; and 3) illustrate the dependence of infection risks between individuals caused by epidemic HIV transmission dynamics.

The results of this investigation are encouraging. This experiment indicates that for an RPT HIV vaccine trial with a sample size in the thousands, VE_S, VE_I and VE_R estimates could be made with considerable accuracy and substantial statistical power. In simulated RPT vaccine trials with a sample size of N=2640, it was observed that the overall accuracy of the VE_S, VE_I and VE_R estimates was quite good. While some slight bias was evident in the estimates, the standard errors and confidence intervals were quite acceptable. Furthermore, the statistical power of the RPT design to detect VE_R was excellent. Consider the simulation results for a vaccine that reduced a recipient's susceptibility by only 25 percent (i.e., VE_S=0.25), but which also reduced the recipient's infectiousness to others by 75 percent if the recipient became infected (i.e., VE_I=0.75). The simulation experiment demonstrated a greater than 90 percent chance that such a vaccine with VE_R of 0.8125 would be considered efficacious by an RPT trial. In contrast, a standard vaccine trial design would have had a 50 percent chance of rejecting such a highly effective vaccine due to its low VE_S value of 0.25.

The simulation parameters selected for this study produced HIV epidemics which were consistent with the conditions which might exist for an HIV vaccine field trial (Barth-Jones 1999). The simulation parameters used in this experiment yielded a modest baseline HIV incidence (about 3 percent annually), a relatively low per-partnership transmission probability (0.092), and a plausible number of trial partners (6.15) in a two year period. Furthermore, the sample size considered here was slightly more than half the sample size of the current Phase III HIV vaccine trials in the United States and the simulated vaccine trials lasted one year less than the U.S. trials are planned to run (Francis 1998, Balter 1998). Under these conservative conditions, excellent statistical power was observed for this experiment, but it should also be noted that this simulation study did not account for the high proportion of missing data that should be anticipated when attempting to contact-trace the sex partners of the infected index cases. However, previous work using a probability model for the statistical power of RPT trials indicates that sufficient statistical power can be obtained, even with substantial missing partner data, if the partner data is randomly missing (Barth-Jones 1999). Further work is needed, though, to investigate the impact of non-random missing partner data on the VE estimates produced by the RPT design.

It was also possible to demonstrate that the infectious disease transmission dynamics produced by the HIVSIM computer model resulted in interdependent outcomes among study participants and thus violated the *iid* assumptions inherent in the use of the Binomial distribution to model infection risks. The correlation of the risks in the vaccinated and placebo treatment arms of the simulated study sites was produced by this interdependence of risks. This interdependence and the associated correlation has been investigated previously through the use of mathematical models, thus allowing proofs regarding the positive nature of this correlation and bounds for the degree of correlation to be established (Donnelly 1993, Helander et al. 1994), but the demonstration of interdependence and correlation using simulation is accessible to a broad audience without recourse to complicated mathematics. Because the correlation of risks in ratio estimates and the associated extra-binomial variation will affect the bias of a ratio estimator (Rice 1995, Collet 1994), the resulting effects of such interdependence on statistical estimates can be complex. Clearly, this is an important area for future

work by epidemiologists, biostatisticians and simulationists. As was seen in this investigation, Monte Carlo simulations involving the explicit simulation of epidemic transmission dynamics are likely to be an important tool in developing a better understanding of this problem and potential analytic solutions (Boily et al. 1996, Chick et al. 1999, Barth-Jones 1999, Desai 1999, Becker 1989).

In conclusion, the results of this Monte Carlo experiment suggest that, using an RPT HIV vaccine trial design, VE_S, VE_I and VE_R estimates can be made with considerable accuracy and substantial statistical power when sufficiently large sample sizes are employed. Given the important role that infectiousness effects might have in halting epidemic HIV transmission, the RPT design warrants further consideration and development as a possible means by which such infectiousness effects might be measured.

ACKNOWLEDGMENTS

Andrew Adams and James Koopman received support from the Center for Disease Control, Alanta, (grant R18/CCR515246-02). James Koopman is also supported by the National Institutes of Health (grant R01-AI45168-01A1). The authors thank Stephen Chick for his insightful and constructive comments on this manuscript.

REFERENCES

Ackerman, E., Elveback, L.R., and J.P. Fox. 1984. *Simulation of infectious Disease Epidemics.* Springfield, Illinois: C.C. Thomas.

Adams, A.L., Barth-Jones, D.C., Chick, S.E., and J.S. Koopman. 1998. Simulations to evaluate HIV vaccine trial designs. *Simulation* 71(4):228-241.

Anderson, R.M., and R.M. May. 1991. *Infectious Diseases of Humans: Dynamics and Control*. Oxford: Oxford University Press.

Balter M. 1999. AIDS now world's fourth biggest killer. *Science*. 284:1101.

Barth-Jones, D.C., Adams, A., Lange, K., Chick, S., and J.S. Koopman. 1998. Retrospective partner trials HIV vaccine study design for measurement of vaccine effects. *Proceedings of the12th World AIDS Conference, Prevention and Epidemiology, Geneva, Switzerland,*; Bolonga, Italy: Monduzzi Editore. pp. 285-289.

Barth-Jones, D.C. 1999. The Retrospective partner trials (RPT) HIV vaccine study design for the measurement of vaccine effects on susceptibility and infectiousness. *Ph.D. Thesis, Epidemiologic Science*, University of Michigan, Ann Arbor, Michigan.

Becker, N.G. 1989. Analysis of infectious disease data. London: Chapman and Hall.

Blackwelder, W.C. 1993. Sample size and power for prospective analysis of relative risk. *Statistics in Medicine* 12:691-698.

Blower, S.M., and A.R. McLean. 1994. Prophylactic vaccines, risk behavior change, and the probability of eradicating HIV in San Francisco. *Science* 265(5177):1451-4.

Boily, M. C. and R. M. Anderson. 1996. Human immunodeficiency virus transmission and the role of other sexually transmitted diseases. Measures of association and study design. *Sexually Transmitted Diseases* 23(4):312-332.

Chick, S.E., Adams, A.L., and J.S. Koopman. 2000. Analysis and simulation of a stochastic, discrete-individual model of STD transmission with partnership concurrency. *Mathematical Biosciences* 166:45-68.

Chick, S.E., Barth-Jones, D.C., and J.S. Koopman. *In Press*. Bias reduction for risk ratio and vaccine effect estimators. *Statistics in Medicine*.

Collett, D. 1994. *Modelling Binary Data*. London: Chapman and Hall.

Datta, S., Halloran, M.E., and I.M. Longini, Jr. 1998. Augmented HIV vaccine trial design for estimating reduction in infectiousness and protective efficacy. *Statistics in Medicine* 17:185-200.

Desai, K.N., Boily, M.C., Masse, B.R., Alary, M., and R.M. Anderson. 1999. Simulation studies of phase III clinical trials to test the efficacy of a candidate HIV-1 vaccine. *Epidemiology and Infection*. 123:65-88.

Donnelly P. 1993. The correlation structure of epidemic models. *Mathematical Biosciences* 117:49-75.

Efron, B., and R.J. Tibshirani. 1993. *An Introduction to the Bootstrap.* New York: Chapman and Hall.

Fox, J.P., Elveback, L., Scott, W., Gatewood, L., and E. Ackerman. 1971. Herd Immunity: basic concept and relevance to public health immunization practices. *American Journal of Epidemiology*. 94(3):179-189.

Francis, D.P., Gregory, T., McElrath, M.J., Belshe, R.B., Gorse, G.J., Migasena, S., Kitayaporn, D., Pitisuttitham, P., Matthews, T., and D.H. Schwartz. 1998. Advancing AIDSVAX to phase 3. Safety, immunogenicity, and plans for phase 3. *AIDS Research and Human Retroviruses.* 14(Suppl 3):S325-331.

Ghani, A.C., Donnelly, C.A., and G.P. Garnett. 1998. Sampling biases and missing data in explorations of sexual partner networks for the spread of sexually transmitted diseases. *Statistics in Medicine*. 17:2079-2097.

Good, P. 1994. *Permutation Tests*. New York: Springer-Verlag.

Haber, M., Longini, I.M. Jr., and M.E. Halloran. 1991. Measures of the effects of vaccination in a randomly

mixing population. *International Journal of Epidemiology* 20(1):300-10.

Helander, M.E., and R. Batta. 1994. A discrete transmission model for HIV. In Kaplan, E.H., and M.I. Brandeu. Eds. *Modeling the AIDS epidemic: Planning, policy and prediction.* New York: Raven Press; pp. 585-611.

Isham, V., and G. Medley. 1996. *Models for Infectious Human Diseases: Their Structure and Relation to Data.* Cambridge: Cambridge University Press.

Jacquez, J.A. 1996. *Compartmental Analysis in Biology and Medicine.* 3rd ed., Biomedware, Ann Arbor, Michigan.

Jewell, N.P. 1986. On the bias of commonly used measures of association for 2 x 2 tables. *Biometrics* 42:351-358.

Leitner, T., Escanilla, D., Franzen, C., Uhlen, M., and J. Albert. 1996. Accurate reconstruction of a known HIV-1 transmission history by phylogenetic tree analysis. *Proceedings of the National Academy of Sciences USA* 93:10864-10869.

Longini, I.M. Jr., Datta, S., and M.E. Halloran. 1996. Measuring vaccine efficacy for both susceptibility to infection and reduction in infectiousness for prophylactic HIV-1 vaccines. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology* 13(5):440-447.

Longini, I.M. Jr., Hudgens M.G., Halloran, M.E., and K.A. Sagatelian 1999. Markov model for measuring vaccine efficacy for both susceptibility to infection and reduction in infectiousness for prophylactic HIV-1 vaccines. *Statistics in Medicine.* 18:53-68.

Mollison, D. 1995. *Epidemic Models: Their Structure and Relation to Data.* Cambridge: Cambridge University Press.

Mooney, C.Z. 1997. *Monte Carlo Simulation. #116: Quantitative Applications in the Social Sciences:* Thousand Oaks, CA: Sage Publications.

O'Neill, R.T. 1988. On sample sizes to estimate the predictive efficacy of a vaccine. *Statistics in Medicine* 7:1279-1288.

Rice, J.A. 1995. *Mathematical Statistics and Data Analysis.* Belmont, California: Duxbury Press.

Rida, W.N. 1996. Assessing the effect of HIV vaccination on infectiousness. *Statistics in Medicine* 15(21-22):2393-404; discussion 2405-2412.

AUTHOR BIOGRAPHIES

DANIEL C. BARTH-JONES is an Assistant Professor at the Center for Healthcare Effectiveness Research and the Department of Internal Medicine at the Wayne State University, School of Medicine. He received his Ph.D. in Epidemiologic Science from the University of Michigan, where he also completed his Master of Public Health degree in General Epidemiology. His primary research interests include vaccinology, the epidemiology of HIV and sexually transmitted diseases, and computer simulation modeling of infectious diseases. His email address is <dbjones@med.wayne.edu>.

ANDREW L. ADAMS is a Software Engineer and the lead architect and developer of the HIVSIM model. He developed the model while working in University of Michigan Department of Epidemiology. He is now leader of the Multicast Engineering Team at NextHop Technologies where he oversees designing, implementing and deploying Internet multicast routing technologies. Adams holds a BSE in Computer Engineering and an MSE in Computer Science, both from the University of Michigan. His research interests include stochastic and agent-based modeling and simulation, and Internet routing and management technologies. His email address is <ala@djinesys.com>.

JAMES S. KOOPMAN is Professor of Epidemiology at the University of Michigan, where he received his medical degree. Koopman also holds a Master's Degree in Public Health. After a residency in pediatrics, he served as an epidemic intelligence service officer with the CDC. His research career is focused on infectious disease epidemiology and the science of transmission system analysis. His email address is <jkoopman@sph.umich.edu>.