

UC Davis

UC Davis Previously Published Works

Title

A discrete-time survival model with random effects for designing and analyzing repeated low-dose challenge experiments

Permalink

<https://escholarship.org/uc/item/88b8j1sw>

Journal

Biostatistics, 16(2)

ISSN

1465-4644

Authors

Kang, Chaeryon
Huang, Ying
Miller, Christopher J

Publication Date

2015-04-01

DOI

10.1093/biostatistics/kxu040

Peer reviewed

A discrete-time survival model with random effects for designing and analyzing repeated low-dose challenge experiments

CHAERYON KANG, YING HUANG*

*Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center,
Seattle, WA 98109, USA*

yhuang@fhcrc.org

CHRISTOPHER J. MILLER

*Center for Comparative Medicine and California National Primate Research Center, University of
California, Davis, CA 95616, USA*

SUMMARY

Repeated low-dose (RLD) challenge designs are important in HIV vaccine research. Current methods for RLD designs rely heavily on an assumption of homogeneous risk of infection among animals, which, upon violation, can lead to invalid inferences and underpowered study designs. We propose to fit a discrete-time survival model with random effects that allows for heterogeneity in the risk of infection among animals and allows for predetermined challenge dose changes over time. Based on this model, we derive likelihood ratio tests and estimators for vaccine efficacy. A two-stage approach is proposed for optimizing the RLD design under cost constraints. Simulation studies demonstrate good finite sample properties of the proposed method and its superior performance compared to existing methods. We illustrate the application of the heterogeneous infection risk model on data from a real simian immunodeficiency virus vaccine study using Rhesus Macaques. The results of our study provide useful guidance for future RLD experimental design.

Keywords: Discrete-time survival model with random effects; Heterogeneous infection risk; HIV vaccine/prevention research; Repeated low-dose challenge experiment; Sample size calculation.

1. INTRODUCTION

Recently, a repeated low-dose (RLD) challenge design has become a standard approach to using non-human primate (NHP) models in HIV vaccine research. Instead of exposing animals to a single high dose of virus to induce infection as in traditional challenge experiments, in RLD experiments animals are repeatedly challenged with a relatively low dose of virus. For example, [Qureshi and others \(2012\)](#) adapted RLD experiments to determine whether the combination of host range mutant adenovirus type-5 and simian immunodeficiency virus (SIV; Ad5 SIVmac239 Gag/Pol/Nef) vaccine has a significant effect on SIV

*To whom correspondence should be addressed.

infection. In this study, animals were randomized to five treatment groups; all animals were repeatedly challenged with the same three escalating doses of SIV viruses regardless of treatment group. RLD experiments have advantages compared with the single high-dose design: they reflect the low probability of HIV transmission in humans more realistically and can provide more statistical power to detect the effects of HIV vaccine (Ellenberger *and others*, 2006; García-Lerma *and others*, 2008; Hudgens and Gilbert, 2009; Hudgens *and others*, 2009; Reynolds *and others*, 2010).

In HIV vaccine research using RLD experiments, it is of major interest to test the effects of the vaccine in preventing HIV infection and to evaluate the magnitude of the vaccine efficacy. Vaccine efficacy is characterized by the vaccine-induced percent reduction in the risk of infection either at each challenge or up to a certain time point during the experiment. Vaccine efficacy has been commonly evaluated using the nonparametric log-rank test and Kaplan–Meier estimator (García-Lerma *and others*, 2008; Reynolds *and others*, 2010; Qureshi *and others*, 2012) or a discrete-time survival model. In an RLD experiment, animals are examined for infection status after each challenge; this produces repeated binary infection outcomes each time. The probability of infection after each challenge can be expressed as a product of Bernoulli trials in the discrete-time survival model.

Current discrete-time survival models to estimate the effects of vaccine in RLD experiments rely heavily on an assumption of homogeneous risk of infection among animals (e.g. García-Lerma *and others*, 2008; Qureshi *and others*, 2012), which, upon violation, can lead to invalid inferences and underpowered study designs. Ignoring heterogeneity among animals can result in underestimated standard errors of the vaccine efficacy estimates and result in confidence intervals with poor coverage (Hudgens and Gilbert, 2009; Moerbeek, 2012). To relax the homogeneity assumption, Hudgens and Gilbert (2009) modeled the transmission probability with a β -distribution assuming independent transmission probabilities across challenges within animals. In this study, we propose to use a discrete-time survival model with random effects to model data from the RLD design, assuming an animal's risks of infection across challenges are independent of each other conditional on random effects. The conditional independence assumption is realistic considering the potential heterogeneity among animals due to biological variation and unobserved covariates. By incorporating random effects, our model flexibly accommodates heterogeneity among animals as well as within-animal dependence with respect to the risk of infection after each challenge.

This article has two goals. The primary goal is to develop a flexible statistical model that can take into account between-animal heterogeneity in RLD experiments while allowing for adjustment of covariates such as the time-dependent challenge doses, as in Qureshi *and others* (2012). In the present study, we propose to fit a discrete-time survival model with a γ -distributed random effect and a complementary log–log (clog–log) link function, which allows for closed forms for the marginal likelihood function and the vaccine efficacy. We estimate model parameters by maximizing the marginal likelihood function and derive asymptotic variance formulae for inferences about vaccine efficacy.

The second goal of this article is to provide guidance on how to design future RLD experiments in terms of the choices for sample size and maximum number of challenges per animal under limited resources. While there is an intensive literature on the design of clinical trials using a single high-dose challenge, design components in RLD experiments have been seldom studied. Because of the complexity of the design, there is no simple analytic formula for calculating sample size and/or maximum number of challenges given desired operational criteria. Previously, Hudgens and Gilbert (2009) used simulation studies to investigate the effects of sample size and maximum number of challenges on statistical power in RLD experiments under the assumption of homogenous risk of infection. In the present study, we propose a two-stage procedure to determine the optimal RLD design under financial constraints allowing for between-animal heterogeneity in the risk of infection.

The remainder of this article is organized as follows. In Section 2, we introduce a heterogeneous infection risk model and an inference about vaccine efficacy. The optimization of the study design for RLD experiments is discussed in Section 3. We investigate performance of the proposed methods through two

intensive simulation studies in Section 4. In Section 5, an application of the proposed risk modeling method to the NHP study described in [Qureshi and others \(2012\)](#) is presented. Finally, we conclude the article with a discussion of our findings and future research topics.

2. A HETEROGENEOUS INFECTION RISK MODEL

2.1 Notation and assumptions

Here we consider an RLD experiment in which each animal is randomly assigned to a treatment (vaccine) or a placebo group, and all animals are repeatedly challenged with the same series of SIVs. Let i be the subject index, Y_i and C_i be time to infection and time to censoring, respectively, and $T_i = Y_i \wedge C_i$ be the time to event outcome variable whichever comes first between Y_i and C_i . An infection indicator is denoted by $\delta_i = \mathbf{1}\{Y_i \leq C_i\}$, where $\mathbf{1}\{A\} = 1$ if A is satisfied, 0 otherwise. Suppose the animals' infection statuses are collected after each challenge. T_i is thus discrete and can be represented as the total number of challenges an animal receives until infected or censored. Let X_i be a $p \times 1$ vector of time-invariant baseline covariates such as the indicator of vaccination status. Let $Z_i(t)$ be a vector of time-variant covariates associated with the challenge dose that animal i receives at time t . For example, when d different levels of challenge doses are applied to each animal, in this article we model $Z_i(t) = \{Z_{i1}(t), Z_{i2}(t), \dots, Z_{id}(t)\}$, where $Z_{i1}(t) = 1$ for all dose levels (reference level), $Z_{ik}(t) = 1$ for the k th dose level, and $Z_{ik}(t) = 0$ otherwise for $k = 2, \dots, d$. Alternatively, $Z_i(t)$ can include polynomial terms of challenge doses at time t if one is interested in continuous dose effects. Let $Z_i^H(t) = \{Z_i(u) : 1 \leq u \leq t\}$ denote the history of Z_i up to time t . We observe n independent identically distributed (i.i.d.) samples, $[O_i = \{T_i, X_i, Z_i^H(T_i), \delta_i\}]_{i=1}^n$. We consider a study design where dose levels are predetermined by protocol for the whole study period and do not vary with individual subjects. That is, $Z_i(t) = Z(t)$ is determined in advance for all individuals under study for t within the planned study period and is conceptually an external time-dependent covariate as described in [Kalbfleisch and Prentice \(2011\)](#).

Let R_i denote the subject-specific random effect. We make the following three assumptions. First, the infection risks across challenges within an animal are assumed to be independent of each other conditional on the random effect and covariates included in the risk model. Second, non-informative censoring is assumed, conditional on the random effect and covariates, which is a reasonable assumption for a well-controlled RLD experiment. Finally, we assume that there is no "memory" of challenge history as commonly assumed in RLD literature ([García-Lerma and others, 2008](#); [Hudgens and Gilbert, 2009](#); [Hudgens and others, 2009](#); [Regoes, 2012](#)). That is, for an animal not yet infected before a challenge, the probability of infection at this particular challenge depends only on the current challenge but not on previous challenges.

2.2 A discrete-time survival model with random effects and a marginal likelihood approach

Let β and α represent vectors of regression parameters for X and Z , for example, the effects of vaccination and challenge doses ($\alpha = (\alpha_1, \dots, \alpha_d)^\top$), respectively. To take into account between-animal heterogeneity under the aforementioned assumptions, we model the risk of infection as follows:

$$\lambda_i\{t|Z_i^H(t), X_i, R_i\} = P\{T_i = t|T_i \geq t, Z_i^H(t), X_i, R_i\} = g^{-1}\{Z_i(t)\alpha + X_i\beta + R_i\}, \quad (2.1)$$

where the random effect R_i follows a specific distribution and g is a link function. Model (2.1) could be extended to include interactions between dose-levels and treatment: $\lambda_i\{t|Z_i^H(t), X_i, R_i\} = g^{-1}\{Z_i(t)\alpha + X_i\beta + \tilde{Z}_i(t)X_i\gamma + R_i\}$, where $\tilde{Z}_i(t) = \{Z_{i2}(t), \dots, Z_{id}(t)\}$, and γ quantifies how the treatment effect changes with dose level. In this study, we use a clog-log link for g and assume that $W_i = \exp(R_i)$ is

independent of other covariates and follows a γ -distribution with mean 1 and variance ν . A γ -distribution with various shapes covers many commonly used exponential-family distributions. Modeling the risk of infection (2.1) with a γ -distributed random effect and a clog-log link function leads to a closed-form expression for the marginal likelihood and the vaccine efficacy, which is computationally efficient. We refer to Conaway (1990), Scheike and Jensen (1997), and Coull and others (2006) for further discussion on the advantages of this combination of the random effects model and link function.

In a discrete-time survival model, given that Z_i is an external covariate, the conditional survival function at time t is $S_i(t|Z_i^H(t), X_i, R_i) \equiv P(T_i > t|Z_i^H(t), X_i, R_i) = P(T_i \neq t|T_i \geq t, Z_i^H(t), X_i, R_i) \times P(T_i \neq t-1|T_i \geq t-1, Z_i^H(t-1), X_i, R_i) \times \cdots \times P(T_i \neq 1|T_i \geq 1, Z_i^H(1), X_i, R_i) = \prod_{j=1}^t \{1 - \lambda_i(j|Z_i^H(j), X_i, R_i)\}$, and the conditional probability of infection at time t is $f_i(t|Z_i^H(t), X_i, R_i) = P(T_i = t|Z_i^H(t), X_i, R_i) = P(T_i = t|T_i \geq t, Z_i^H(t), X_i, R_i) \times P(T_i \neq t-1|T_i \geq t-1, Z_i^H(t-1), X_i, R_i) \times \cdots \times P(T_i \neq 1|T_i \geq 1, Z_i^H(1), X_i, R_i) = \lambda_i(t|Z_i^H(t), X_i, R_i)S_i(t-1|Z_i^H(t-1), X_i, R_i)$. The marginal likelihood function for n subjects indexed by $\theta = (\alpha, \beta, \nu)$ is given by

$$L(\theta; O) = \prod_{i=1}^n \left\{ \int S_i(T_i|Z_i^H(T_i), X_i, R_i; \theta) dG(R) \right\}^{1-\delta_i} \left\{ \int f_i(T_i|Z_i^H(T_i), X_i, R_i; \theta) dG(R) \right\}^{\delta_i}, \quad (2.2)$$

where G is the cumulative distribution function (CDF) of R . Let $\eta_i(j) = Z_i(j)\alpha + X_i\beta$ denote the linear predictor at the time of challenge j . As shown in supplementary material available at *Biostatistics* online, Section S1,

$$\int S_i(T_i|Z_i^H(T_i), X_i, R_i; \theta) dG(R) = \{\nu A_i(T_i) + 1\}^{-1/\nu},$$

$$\int f_i(T_i|Z_i^H(T_i), X_i, R_i; \theta) dG(R) = \{\nu A_i(T_i - 1) + 1\}^{-1/\nu} - \{\nu A_i(T_i) + 1\}^{-1/\nu},$$

where $A_i(t) = \sum_{j=1}^t \exp\{Z_i(j)\alpha + X_i\beta\}$ assuming $A_i(0) = 0$. Therefore, we have a closed-form formula for the marginal log-likelihood function:

$$\ell(\theta; O) = n\mathbb{P}_n[(1 - \delta) \log(\nu A(T) + 1)^{-1/\nu} + \delta \log\{(\nu A(T - 1) + 1)^{-1/\nu} - (\nu A(T) + 1)^{-1/\nu}\}],$$

where \mathbb{P}_n is the standard empirical measure for $O = \{O_i\}_{i=1}^n$.

2.3 Estimation and inference

We estimate parameters in (2.1) and variance ν by maximizing the log likelihood (2.2), for example, through Fisher scoring using iteratively reweighted least squares. The variability of the random effects can be expressed more intuitively by an intraclass correlation, ρ , between underlying continuous responses. In particular, let $B_i(t)$ be a binary outcome indicating infection status for animal i after the t th challenge assuming no event occurs before t . Let $C_i(t)$ be the underlying continuous latent outcome such that $B_i(t) = 1$ if $C_i(t) > 0$ and $B_i(t) = 0$ otherwise. Suppose $C_i(t) = \eta_i(t) + R_i + \epsilon_i(t)$, where $-\epsilon_i(t)$, the individual error term, has a reverse extreme value distribution with the CDF $H\{\epsilon_i(t)\} = \exp[-\exp\{-\epsilon_i(t)\}]$ and variance $\pi^2/6$. Under the model with a γ -distributed random effect and a clog-log link, the intrasubject correlation coefficient for the underlying continuous outcome equals $\rho = \nu/(\nu + \pi^2/6)$. Equivalently, we have $\nu = \pi^2\rho/\{6(1 - \rho)\}$. For more details, we refer to Coull and others (2006) and Rodriguez and Elo (2003).

The null hypothesis of no between-animal heterogeneity, $\nu = 0$, is equivalent to the null hypothesis of no random effects, $R = 0$, for all animals. Under the null hypothesis,

$$L(\theta|O, R = 0) = \prod_{i=1}^n [\exp\{-A_i(T_i)\}]^{1-\delta_i} [\exp\{-A_i(T_i - 1)\} - \exp\{-A_i(T_i)\}]^{\delta_i}, \quad (2.3)$$

which is a likelihood function assuming independent risk of infection across challenges. As shown in supplementary material available at *Biostatistics* online, Section S1, (2.2) converges to (2.3) as $\nu \rightarrow 0$. Let $\Theta^\nu = \{\theta = (\alpha, \beta, \nu) : \alpha \in \mathbb{R}^q, \beta \in \mathbb{R}, \nu \in (0, \infty)\}$ and $\Theta_0^\nu = \{(\alpha, \beta, \nu) : \alpha \in \mathbb{R}^q, \beta \in \mathbb{R}, \nu = 0\}$ denote parameter spaces under the alternative and null hypotheses of zero between-animal heterogeneity, respectively. We conduct the likelihood ratio test (LRT) to reject the null hypothesis for a large value of the LR statistic: $\Lambda_\nu = \sup L\{\nu|\cdot, \theta \in \Theta_0^\nu\} / \sup L\{\nu|\cdot, \theta \in \Theta^\nu\}$. Under the null hypothesis of $\nu = 0$, the value of ν lies on the boundary of the parameter space, $\nu \in [0, \infty)$, such that Λ_ν converges to a mixture of χ^2 distributions $0.5\chi_{1-\alpha}^2(0) + 0.5\chi_{1-\alpha}^2(1) = 0.5\chi_{1-\alpha}^2(1)$, where $\chi_{1-\alpha}^2(r)$ is the $100(1 - \alpha)$ th percentile of a χ^2 distribution with r degrees of freedom (d.f.), as presented in [Self and Liang \(1987\)](#), [Goldman and Whelan \(2000\)](#), and [Hudgens and Gilbert \(2009\)](#). We reject $H_0 : \nu = 0$ if $-2 \ln \hat{\Lambda}_\nu > 0.5\chi_{1-\alpha}^2(1)$.

The test for the effect of vaccine is equivalent to the test of the null hypothesis $H_0 : \beta = 0$. Let $\Theta^\beta = \{(\alpha, \beta, \nu) : \alpha \in \mathbb{R}^q, \beta \in \mathbb{R}, \nu \in [0, \infty)\}$ and $\Theta_0^\beta = \{(\alpha, \beta, \nu) : \beta = 0, \alpha \in \mathbb{R}^q, \nu \in [0, \infty)\}$ denote parameter spaces under the alternative and null hypotheses of a zero vaccine effect, respectively. We reject the null hypothesis if $-2 \ln \hat{\Lambda}_\beta > \chi_{1-\alpha}^2(r)$, where $\Lambda_\beta = \sup L\{\beta|\cdot, \theta \in \Theta_0^\beta\} / \sup L\{\beta|\cdot, \theta \in \Theta^\beta\}$, and r is the difference in the number of parameters between the two-nested models.

2.4 Estimation of vaccine efficacy

[Hudgens and Gilbert \(2009\)](#) defined two types of vaccine efficacy. The first is vaccine efficacy for preventing infection before or at the time of challenge t :

$$\text{VE}(t) = 1 - \frac{\text{probability of infection before or at time } t \text{ for the vaccine group}}{\text{probability of infection before or at time } t \text{ for the placebo group}},$$

which is the relative reduction in the risk of infection before or at time t for the vaccine group compared to the placebo group. $\text{VE}(t) > 0$ indicates that the vaccine is effective in reducing the risk of infection before or at time t , whereas $\text{VE}(t) \leq 0$ indicates that the vaccine is not effective or has a negative effect. Under the heterogeneous infection risk model described in Section 2.2, $\text{VE}(t) = 1 - \left(1 - \left[\nu \sum_{j=1}^t \exp\{Z_i(j)\alpha + \beta\} + 1\right]^{-1/\nu}\right) / \left(1 - \left[\nu \sum_{j=1}^t \exp\{Z_i(j)\alpha\} + 1\right]^{-1/\nu}\right)$. The second type of vaccine efficacy is the perchallenge vaccine efficacy, defined as the relative reduction in the risk of infection caused by vaccination at a particular challenge, conditional on non-infection before the challenge. Perchallenge VE at dose-level k under the heterogeneous infection risk model equals $\text{VE}_k = 1 - (1 - \{\nu \exp(Z_k\alpha + \beta) + 1\}^{-1/\nu}) / (1 - \{\nu \exp(Z_k\alpha) + 1\}^{-1/\nu})$, where Z_k is a vector of variables of length d for the k th dose level with the ℓ th element being $\mathbf{1}\{\ell = k\}$, $\ell = 2, \dots, d$ and $Z_1 = 1$ (reference level). VE_k allows characterization of the vaccine's effect at a specific level of exposure, whereas $\text{VE}(t)$ represents a vaccine effect integrated over multiple levels of exposures. Let $\hat{\text{VE}}(t)$ and $\hat{\text{VE}}_k$ be MLE of $\text{VE}(t)$ and VE_k . The covariance matrices of $\hat{\text{VE}}(t)$ and $\hat{\text{VE}}_k$ can be calculated using the Delta method as given in supplementary material available at *Biostatistics* online, Section S1.

3. OPTIMIZATION OF THE DESIGN OF RLD EXPERIMENTS UNDER COST CONSTRAINTS

In practice, we have limited resources for conducting RLD experiments. It is of interest to optimize the study design such that a desired operational criterion, for example, precision of VE estimators or power of the study, can be maximized. As pointed out by Moerbeek (2012) and Zhang and Ahn (2011) for the cluster randomized design with discrete survival outcomes, no analytical formula is available yet for calculating sample size and power; Monte-Carlo simulation studies are commonly used to determine design components. The results of the simulation study in Hudgens and Gilbert (2009) demonstrated that the statistical power for the test of vaccine efficacy increases with larger sample size, larger maximum number of challenges per animal, and higher risk of infection per exposure in the control group. Optimizing an RLD design typically requires excessive simulations under various combinations of sample size and number of challenges.

Motivated by the study in Zhang and Ahn (2011), we propose a computationally efficient two-stage approach to identify the optimal pair of sample size and maximum number of challenges, denoted by (n^*, m^*) , in order to maximize the operational criteria of interest under financial constraints. We illustrate how the proposed approach can be used to guide the design of an RLD experiment to optimize the efficiency of estimating perchallenge vaccine efficacy, VE_k . The same strategy applies to other criteria such as $VE(t)$ and study power.

Suppose the financial costs of adding an animal and adding a challenge per animal are C_1 and C_2 , respectively, and the maximum budget C for the experiment is fixed. Suppose d unique levels of escalating challenge doses will be applied to each animal as in Qureshi and others (2012), and each dose level is applied to an animal for m times ($m \geq 2$). Let q_j denote the probability of an animal receiving the j th challenge for $j = 1, \dots, d \times m$. The average cost to challenge n animals for up to m times at each dose level equals $nC_1 + n \sum_{j=1}^{d \times m} q_j C_2 \leq C$. Define $n(m)$ as the largest n for a given m satisfying this constraint: $n(m) \equiv \lfloor C / (C_1 + \sum_{j=1}^{d \times m} q_j C_2) \rfloor$, where $\lfloor x \rfloor$ denotes the largest integer not greater than x . Since n and m are not independent conditional on C , seeking the best pair (n^*, m^*) that minimizes the variance for estimators of VE_k is not straightforward in practice, particularly for a complex study design like an RLD experiment.

Let $S_{(n,m)} = \{ \{n(m), m\} : \text{form } m \in M = \{1, \dots, m_{\max}\} \}$ for a prespecified m_{\max} . We propose to use a two-stage approach to find (n^*, m^*) . In the first stage, we conduct a simulation study with a large sample size (N) to estimate $\text{Var}(\hat{VE}_k)$ for each $m \in M$, denoted by $\hat{\text{Var}}(\hat{VE}_k)_{\{N,m\}}$. In the second stage, we find the $\{n(m), m\} \in S_{(n,m)}$ that minimizes the variance estimate of \hat{VE}_k . In particular, let $\hat{\text{Var}}(\hat{VE}_k)_{\{n(m),m\}}$ denote the variance estimate based on data with $\{n(m), m\}$. We approximate $\hat{\text{Var}}(\hat{VE}_k)_{\{n(m),m\}}$ with $\hat{\text{Var}}(\hat{VE}_k)_{\{N,m\}} \times N/n(m)$. Using $m = 2$ as the reference design, we evaluate the efficiency at $\{n(m), m\}$ relative to $\{n(2), 2\}$ by

$$\text{RE}\{n(m), m\} = \frac{\hat{\text{Var}}(\hat{VE}_k)_{\{n(2),2\}}}{\hat{\text{Var}}(\hat{VE}_k)_{\{n(m),m\}}}. \quad (3.1)$$

We compute $(n^*, m^*) \equiv \arg \max_{\{n(m),m\} \in S_{(n,m)}} \text{RE}\{n(m), m\}$. The resultant (n^*, m^*) would achieve the best efficiency for estimating VE_k under a fixed total cost. Compared with the typical Monte-Carlo simulation studies that evaluate variances of \hat{VE}_k estimators for every n, m combinations, the computational burden reduces from $O(n) \times O(m)$ to $O(m)$ using the proposed two-stage approach.

4. SIMULATION STUDIES

We illustrate the proposed heterogeneous infection risk model and the two-stage approach to optimizing RLD design with two intensive simulation studies. In the first simulation study, we compare the discrete-time survival model with random effects (hereafter ‘‘heterogeneous model’’) with a discrete-time survival

model assuming independence in the risk of infection across challenges within animals (“homogeneous model”), with respect to statistical power, Type I error rate and the precision of parameter estimates. In the second simulation study, we explore the best pair of sample size and maximum number of challenges per animal for various levels of vaccine efficacy and between-animal heterogeneity under a fixed total cost.

For both simulation studies, we considered RLD studies in which animals are 1:1 randomized to a vaccine group and a placebo group and challenged with the same three increasing levels of challenge doses ($d = 3$) to mimic the setting in [Qureshi and others \(2012\)](#). Each animal was allowed up to $d \times m$ challenges. Let X_i be the indicator of assignment to the vaccine group and $Z_i(j) = \{Z_{i1}(j), Z_{i2}(j), Z_{i3}(j)\}$ as defined in Section 2.1. Let $R_i = \log(U_i)$ be a random effect with U_i generated from a γ -distribution with mean 1 and variance ν . The conditional probability of infection at each challenge was modeled as in (2.1) with the clog–log link function. We observed the binary outcome $Y_i^*(j) \sim \text{Bernoulli}(\lambda_i(j|Z_i^H(j), X_i, R_i))$, $j = 1, \dots, d \times m$, the survival time $T_i \equiv \min\{j \in [1, d \times m] : Y_i^*(j) = 1\}$, and the infection indicator $\delta_i = \mathbf{1}\{T_i \leq d \times m\}$. We maximized the likelihood functions in (2.2) and (2.3) using the `optim` function in the R package ([R Core Team, 2012](#)).

4.1 Simulation Study 1

4.1.1 Set-up. In the first simulation study, each animal was allowed up to 5 challenges at each dose level, with 15 the maximum number of challenges for each animal. Data were generated with different ν values for zero ($\nu = \rho = 0$), weak ($\nu = 0.89$ or $\rho = 0.35$), moderate ($\nu = 2.01$ or $\rho = 0.55$), or strong ($\nu = 4.93$ or $\rho = 0.75$) within-animal dependence. We set $\beta = -1$ and $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ corresponding to perchallenge infection probabilities among placebo recipients of 0.02, 0.16, and 0.39 at dose levels 1, 2, and 3, respectively. Sample sizes $n = 40, 200, \text{ or } 1000$ were explored. For each simulation scenario, 1000 Monte-Carlo data were generated.

4.1.2 Results. Without within-animal dependence ($\nu = 0$), average estimates for model parameters and for perchallenge VEs using the heterogeneous model were comparable with those estimated using the homogeneous model, with slightly larger variances. In the presence of weak within-animal dependence ($\nu = 0.89$), the homogeneous model produced highly biased estimates while the heterogeneous model produced unbiased estimates even when sample size was small (Table 1). In addition, perchallenge VE tended to be underestimated by fitting the homogeneous model with much worse coverage rates, while the heterogeneous model produced unbiased estimates with coverage rates close to the target nominal level (Table 2). Results for the settings with moderate or strong within-animal dependence, presented in supplementary material available at *Biostatistics* online, Tables S1 and S2, showed the apparent superiority of the heterogeneous model. Supplementary material available at *Biostatistics* online, Figure S1, shows the estimate of $\text{VE}(t)$ versus t . The homogeneous model underestimated $\text{VE}(t)$, particularly at the early stages of the study, while the estimated $\text{VE}(t)$ curve using the heterogeneous model was very close to the true $\text{VE}(t)$ curve.

For settings with $\nu > 0$, LRT testing for vaccine effects using the homogenous model had elevated Type I error rates, while LRT using the heterogeneous model and the log-rank test by contrast had Type I error rates close to the target nominal level (Table 3). LRT using the heterogeneous model produced the highest power; the gain in power by accounting for within-animal dependence increased as ν increased. In general, power for testing vaccine effects decreases as ν increases. Consequently, a much larger sample size will be required to detect a weak vaccine effect in the presence of moderate or strong within-animal dependence. Supplementary material available at *Biostatistics* online, Table S3, shows the results of LRT for testing heterogeneity ($H_0 : \nu = 0$) based on the heterogeneous model. LRT had well-controlled Type I

Table 1. Results of simulation study 1 to compare the homogeneous infection risk model (Homogeneous) and heterogeneous infection risk model (Heterogeneous) at the sample sizes $n = 40, 200, \text{ or } 1000$ under the following two cases: data were generated under the independent risk of infection ($\nu = 0$), and data were generated under the weak strength of dependence in the risk of infection ($\nu = 0.89$) across challenges within each animal. Monte-Carlo mean (mean) and standard deviation (MCSD) of estimates and 95% confidence interval coverage rates (%) based on the normal approximation of estimates (95% CR) using 1000 simulations are reported

Fitting model		Data without within-animal dependence ($\nu = 0$)						Data with within-animal dependence ($\nu = 0.89$ ($\rho = 0.35$))							
		True value	Homogeneous			Heterogeneous			True value	Homogeneous			Heterogeneous		
Sample size	40		200	1000	40	200	1000	40		200	1000	40	200	1000	
α_1	Mean	-3.90	-5.33	-3.95	-3.90	-5.27	-3.93	-3.89	-3.89	-5.20	-4.07	-4.03	-5.02	-3.94	-3.90
	MCSD		4.53	0.30	0.12	4.42	0.30	0.12		4.05	0.31	0.13	3.89	0.33	0.14
	95% CR		88.7	95.3	96.0	88.4	95.8	96.0		92.1	94.4	84.7	92.2	95.2	95.8
α_2	Mean	2.16	3.57	2.20	2.16	3.59	2.23	2.17	2.22	3.09	2.01	1.96	3.34	2.28	2.22
	MCSD		4.54	0.31	0.13	4.42	0.32	0.13		4.05	0.34	0.14	3.87	0.36	0.15
	95% CR		89.6	95.5	96.8	89.7	96.2	96.8		86.8	84.2	48.2	91.4	94.5	95.0
α_3	Mean	3.20	4.69	3.25	3.20	4.83	3.36	3.25	3.42	3.78	2.64	2.59	4.62	3.46	3.41
	MCSD		4.54	0.31	0.13	4.42	0.36	0.14		4.07	0.33	0.14	3.96	0.51	0.22
	95% CR		89.5	95.9	95.5	91.3	97.1	98.1		70.5	30.1	0.0	93.3	95.8	95.0
β	Mean	-1.00	-1.05	-1.00	-1.00	-1.14	-1.06	-1.03	-1.00	-0.69	-0.70	-0.68	-1.03	-1.03	-1.00
	MCSD		0.39	0.17	0.07	0.44	0.19	0.08		0.42	0.18	0.08	0.66	0.28	0.12
	95% CR		93.8	94.0	94.5	97.9	98.1	97.3		84.1	56.5	1.3	94.8	95.2	95.8
ν	Mean	0.00				0.13	0.08	0.04	0.89				0.94	0.86	0.87
	MCSD					0.21	0.12	0.05					0.96	0.44	0.19

Mean = $\bar{\hat{\theta}} = \sum_{m=1}^M \hat{\theta}_m / M$, MCSD = $\sqrt{\sum_{m=1}^M (\hat{\theta}_m - \bar{\hat{\theta}})^2 / (M - 1)}$, where $\hat{\theta}_m$ denotes estimates from the m th simulation for $M = 1000$.

Table 2. Results of simulation Study 1 to compare the homogeneous infection risk model (Homogeneous) and heterogeneous infection risk model (Heterogeneous) at the sample sizes $n = 40, 200, \text{ or } 1000$ under the following two cases: data were generated under the independent risk of infection ($\nu = 0$), and data were generated under the weak strength of dependence in the risk of infection ($\nu = 0.89$) across challenges within each animal. Monte-Carlo mean (mean) and standard deviation (MCSD) of perchallenge VE estimates at each dose level ($VE_1, VE_2, \text{ and } VE_3$) and 95% confidence interval coverage rates (%) based on the normal approximation of perchallenge VE (95% CR) and normal approximation of the logit-transformed perchallenge VE (95% CR (L)) using 1000 simulations are reported

Fitting model		Data without within-animal dependence ($\nu = 0$)							Data with within-animal dependence ($\nu = 0.89$ ($\rho = 0.35$))						
		Sample size	True value	Homogeneous			Heterogeneous			True value	Homogeneous			Heterogeneous	
40	200			1000	40	200	1000	40	200		1000	40	200	1000	
VE_1	Mean	0.63	0.62	0.63	0.63	0.65	0.64	0.64	0.63	0.45	0.49	0.49	0.56	0.62	0.62
	MCSD		0.15	0.06	0.03	0.15	0.06	0.03		0.23	0.09	0.04	0.30	0.10	0.04
	95% CR		90.3	94.3	94.6	92.5	95.0	96.3		97.2	73.1	2.6	92.4	93.8	95.7
	95% CR (L)		94.5	95.0	94.9	98.6	97.1	96.8		99.1	72.7	2.6	97.8	96.5	96.4
VE_2	Mean	0.61	0.61	0.61	0.61	0.63	0.62	0.62	0.59	0.44	0.48	0.47	0.53	0.59	0.59
	MCSD		0.15	0.06	0.03	0.15	0.06	0.03		0.22	0.09	0.04	0.27	0.10	0.04
	95% CR		90.6	94.3	94.9	92.3	94.8	96.4		96.7	81.5	9.6	92.4	94.5	95.8
	95% CR (L)		94.7	95.0	95.1	98.3	96.4	96.7		98.4	83.6	9.1	96.7	96.2	96.1
VE_3	Mean	0.57	0.56	0.57	0.57	0.57	0.58	0.58	0.52	0.43	0.47	0.46	0.45	0.51	0.52
	MCSD		0.14	0.06	0.03	0.14	0.06	0.03		0.22	0.09	0.04	0.23	0.09	0.04
	95% CR		91.3	93.8	94.7	90.7	94.4	95.4		95.2	93.7	66.9	93.1	94.5	95.4
	95% CR (L)		94.4	95.2	94.5	94.7	95.1	95.8		96.6	98.4	67.6	96.2	95.8	95.5

Mean = $\bar{\hat{\theta}} = \sum_{m=1}^M \hat{\theta}_m / M$, MCSD = $\sqrt{\sum_{m=1}^M (\hat{\theta}_m - \bar{\hat{\theta}})^2 / (M - 1)}$, where $\hat{\theta}_m$ denotes estimates from the m th simulation for $M = 1000$.

Table 3. Results of simulation Study 1 to compare the Type I error rate and power of the tests for the vaccine effects at different strength of within-animal dependence ($\nu = 0, 0.89, 2.01, \text{ or } 4.93$) and sample sizes ($n = 40, 200, \text{ or } 1000$). Three tests are compared: The likelihood ratio test (LRT) using the homogeneous infection risk model (LRT-Homo); LRT using the heterogeneous infection risk model (LRT-Hetero); and the log-rank test (Log-rank). Rejection rates (%) over 1000 replications using $\chi_{0.95}^2$ (d.f. = 1) distribution under the null ($\beta = 0$) and the alternative ($\beta = -1$) hypotheses are reported for the Type I error rate and power, respectively

		$\nu = 0$			$\nu = 0.89$			$\nu = 2.01$			$\nu = 4.93$		
		LRT		Log-rank	LRT		Log-rank	LRT		Log-rank	LRT		Log-rank
	Sample size	Homo	Hetero	Log-rank	Homo	Hetero	Log-rank	Homo	Hetero	Log-rank	Homo	Hetero	Log-rank
Type I error rate	40	7.2	6.8	7.2	6.7	6.1	5.1	6.8	7.4	5.4	6.4	7.5	5.0
	200	5.9	5.4	5.9	5.8	5.3	4.5	6.9	5.7	5.3	7.1	4.8	5.5
	1000	5.6	5.2	5.2	5.9	4.9	5.1	5.7	5.6	4.5	6.6	5.3	5.4
Power	40	82.6	81.2	81.3	44.0	45.3	41.4	24.6	29.0	22.3	14.9	18.7	12.5
	200	100.0	100.0	100.0	98.3	98.8	98.3	79.9	87.3	79.2	39.2	57.5	37.9
	1000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	95.6	99.8	95.6

error rates, and the power of the test increased with the sample size and the magnitude of the within-animal dependence.

More results of the simulation study investigating the performance of the heterogeneous model are presented in supplementary material available at *Biostatistics* online, Tables S4–S9 in Sections S2 and S3. A similar pattern comparing heterogeneous and homogenous models was observed in the extended model that includes interactions between the dose levels and vaccine assignment (supplementary material available at *Biostatistics* online, Tables S7–S9). We also investigated the robustness of the proposed model under risk model misspecification. Results are presented in supplementary material available at *Biostatistics* online, Tables S10–S11 in Section S4. In general, the proposed heterogeneous model has a compatible or slightly better performance compared with the homogeneous model or the log-rank test under a moderate degree of working model misspecification.

4.2 Simulation Study 2

4.2.1 Set-up. We considered $\nu = 0.89, 2.01, \text{ or } 4.93$ corresponding to $\rho = 0.35, 0.55, \text{ or } 0.75$, respectively. Values for $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ were chosen such that the perchallenge probabilities of infection in the placebo group for $\nu = 0.89, 2.01, \text{ and } 4.93$, equal 0.1, 0.09, and 0.08, respectively, at dose level 1, 0.18, 0.17, and 0.14 at dose level 2, and 0.27, 0.24, and 0.19 at dose level 3. Various values for β corresponding to $VE_1 = 0.2, 0.5, \text{ or } 0.7$ were considered, where VE_1 is perchallenge vaccine efficacy at the lowest level dose. We assumed that the cost for adding an animal is 50 times the cost for an extra challenge per animal and set the maximum budget to be $C = 300 \times C_1$. A range of the maximum number of challenges per dose level, $M = \{m : m = 2, 3, \dots, 15\}$, were considered to allow for at least two challenges at each dose level.

4.2.2 Results. Our goal was to identify the optimal pair of sample size and maximum number of challenges per animal to achieve the best efficiency in estimating VE_1 given budget constraints. In the first stage, we estimated the variance of \hat{VE}_1 for each $\{n(m), m\} \in S_{(n,m)}$ using 10 simulated datasets, each containing $N = 5000$ samples. In the second stage, we calculated $RE\{n(m), m\}$, the efficiency of \hat{VE}_1 for $\{n(m), m\}$ relative to $\{n(2), 2\}$ by (3.1) and averaged $RE\{n(m), m\}$ over the 10 simulated datasets to obtain a smooth estimate. We then identified the best pair of $\{n(m), m\}$ that achieved the greatest average relative efficiency compared with $\{n(2), 2\}$. Figure 1 displays $RE\{n(m), m\}$ for all $\{n(m), m\} \in S_{(n(m),m)}$ and for the selected (n^*, m^*) by the proposed approach. The efficiency gain of the selected pair compared with the reference pair $\{n(2), 2\}$ ranges from 2 to 78% for varying ρ and VE_1 . Under fixed within-animal dependence, with the increase in vaccine efficacy, the maximum efficiency was achieved at a larger number of challenges and thus a smaller sample size. Under fixed vaccine efficacy, with an increase in within-animal dependence, the maximum efficiency was achieved at a larger sample size and thus a smaller number of challenges. Comparisons of the estimated $RE\{n(m), m\}$ based on the two-stage method and those obtained from the standard Monte-Carlo simulations for $VE_1 = 0.5$ and $\rho = 0.35$ or 0.55 are given in supplementary material available at *Biostatistics* online, Figure S2. The pattern of estimated $RE\{n(m), m\}$ versus m was similar between the two.

5. DATA ANALYSIS

In the NHP study described by [Qureshi and others \(2012\)](#), 43 adult male Rhesus Macaca mulatta (RM) were randomized to five treatment groups: Ad5 Vx-SIV ($n = 9$), Ad5 Vx-empty (8), Vx-SIV (9), Vx-empty (9), and naive control (8). All macaques were challenged by a series of escalating penile exposures including three levels of SIVmac 251 viruses: 10^3 TCID₅₀ of virus 10 times weekly (the lowest level dose),

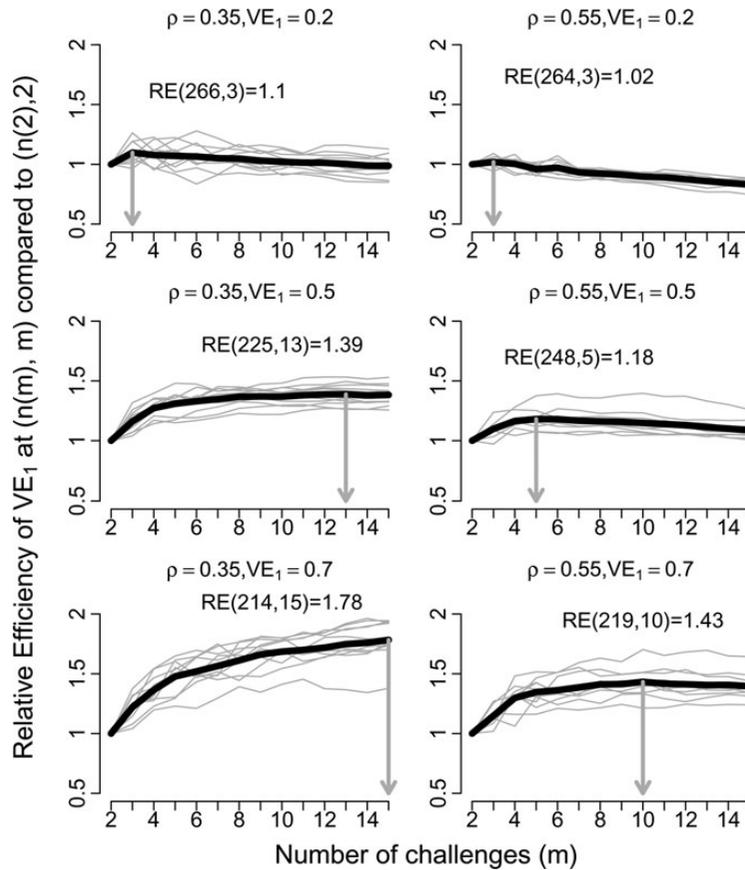


Fig. 1. The relative efficiency (RE) of the VE_1 estimator for each pair of $\{n(m), m\}$ compared with $\{n(2), 2\}$. $RE\{n(m), n\}$ is defined by $RE\{n(m), m\} = \hat{\text{var}}(\hat{VE}_1)_{\{n(2), 2\}} / \hat{\text{var}}(\hat{VE}_1)_{\{n(m), m\}}$. Intraclass correlation $\rho = (0.35, 0.55)$ and perchallenge vaccine efficacy at the first dose ($VE_1 = (0.2, 0.5, 0.7)$) are explored. The selected pair of $\{n(m), m\}$ by the proposed approach is indicated on the X -axis, and $RE\{n(m), m\}$ is shown. Black line and gray lines represent the average $RE\{n(m), m\}$ and $RE\{n(m), m\}$ over 10 replications. The sample size of the data for the first-stage simulation $N = 5000$.

followed by 10^4TCID_{50} of virus 10 times weekly (the middle level dose), and finally 10^5TCID_{50} of virus twice a day (the highest level dose). Animals were evaluated for infection status after each challenge. Time to infection was defined as the first week in which $\log_{10} \text{vRNA}$ copies are > 2 and stay > 2 for subsequent measurements within 4 weeks. The process of immunization can be found in detail in [Qureshi and others \(2012\)](#).

Previously, [Qureshi and others \(2012\)](#) fit the discrete-time survival model to the data assuming homogeneous and constant risk of infection. In this section, we reanalyze the data allowing for possible heterogeneity in the risk of infection across animals and allowing for adjustment of changing dose levels over time. Results for the average risk of infection across different time periods, as presented in [Qureshi and others \(2012\)](#), suggest possible differences in vaccine efficacy across challenge dose levels, particularly between the highest and the other two dose levels. To allow for some flexibility in modeling VE as a function of the challenging dose, we include the interaction term between the highest dose level

and treatment assignment in the model of infection risk, in addition to main effects for each dose level and treatment: $\lambda_i(t|Z_i^H(t), X_i, R_i) = g^{-1}(Z_i(t)\alpha + X_i\beta + \mathbf{1}\{t \in \text{the highest dose level}\}X_i\gamma + R_i)$. A similar interaction model was considered in analyzing an HIV vaccine trial in [Robb and others \(2012\)](#). We fit both the homogeneous and the heterogeneous infection risk models. A goodness-of-fit test ([Pan and Lin, 2005](#)) for the proposed heterogeneous model did not show any significant deviation from the model assumptions (supplementary material available at *Biostatistics* online, Figure S3 in Section S6).

Using the heterogeneous model, the LRT did not find any significant within-animal dependence. However, the power to detect small heterogeneity in this study with 43 animals is very limited. [Regoes \(2012\)](#) studied heterogeneous susceptibility in eight datasets for HIV using an RLD design and failed to find a statistically significant heterogeneous susceptibility except for the data from Letvin 11 stm which had the largest sample size (43 animals for each of two vaccine groups). The results in [Regoes \(2012\)](#) and those in our analysis suggest that larger sample size is important for detecting heterogeneity in the risk of infection. Given the limited power of the test in practice, fitting a heterogeneous model to the data allowing for potential heterogeneity across animals can be useful to complement and/or confirm the results of the simpler homogeneous model, even in the presence of nonsignificant testing results.

There were no statistically significant differences identified in the risk of infection among the five vaccine groups based on the LRT (p -value = 0.202 and 0.202 for the homogeneous and the heterogeneous models, respectively). Based on both the homogeneous and heterogeneous models, estimated perchallenge VE comparing composite groups of particular interest as defined in [Qureshi and others \(2012\)](#) are reported along with 95% confidence intervals (Table 4). The negative perchallenge vaccine efficacy estimate for Ad5-Vx-SIV vaccine against Ad5-Vx-empty or other SIV-negative vaccines in the lowest and the middle dose levels suggests a possibly greater risk of infection in the Ad5 seropositive animals immunized with the Ad5 SIV. This result is consistent with [Qureshi and others \(2012\)](#). More investigation through larger studies, however, is required to confirm this observation. Estimated perchallenge probabilities of infection are given in supplementary material available at *Biostatistics* online, Tables S12–S14 in Section S6. Estimation of perchallenge infection risk and vaccine efficacy for each specific dose level was not achievable in earlier analyses that assumed constant infection probability and VE across challenges and animals. These measures provide valuable information to biologists for planning future RLD experiments.

6. DISCUSSION

In this article, we propose to fit a discrete-time survival model with random effects to take into account the potential heterogeneity among animals and the resulting within-animal dependence arising in RLD experiments. Simulation studies demonstrate that in the presence of heterogeneity, the homogeneous model ignoring within-animal dependence can have an inflated Type I error rate, loss of power for testing the vaccine's effect, biased estimates of vaccine efficacy, and low coverage rates. These problems were resolved by fitting a heterogeneous infection risk model. In the absence of between-animal heterogeneity, the heterogeneous infection risk model produces results comparable with the homogeneous infection risk model. We also propose a two-stage approach to determine the optimal balance between the number of animals and the number of challenges, in order to maximize the efficiency of VE estimates under limited resources. The usefulness of the method in designing an RLD experiment is demonstrated through simulation studies.

The heterogeneous infection risk model is more plausible from a biological point of view compared to a homogeneous model. In practice, however, most NHP studies on HIV vaccines have a small sample size with weak vaccine efficacy, which can lead to insufficient power to detect heterogeneity, as discussed in [Regoes \(2012\)](#). Nevertheless, heterogeneity among animals is commonly believed to exist. Our methods address the need to assess and accommodate the existence of heterogeneity in risk modeling, even when the study may not be powerful enough to lead to significant test results.

Table 4. Perchallenge VE (VE_k) estimates (Estimate) along with standard error estimates (SE) and 95% confidence intervals (95% CI) obtained by the homogeneous and heterogeneous infection risk models including the interaction between the highest dose level and vaccine group in the NHP data (Qureshi and others, 2012) are reported. Confidence interval is built using normal approximation of VE_k

Model	Comparison group	The lowest level virus (10^3TCID_{50})			The middle level virus (10^4TCID_{50})			The highest level virus (10^5TCID_{50})		
		Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
Homogeneous	Ad5 Vx-SIV versus Ad5 Vx-empty	-3.58	5.12	(-13.61, 6.45)	-3.49	4.96	(-13.21, 6.24)	0.65	0.33	(0.00, 1.30)
	Ad5 Vx-SIV + Vx-SIV versus All others	-0.46	0.67	(-1.78, 0.85)	-0.45	0.65	(-1.72, 0.82)	0.29	0.39	(-0.48, 1.05)
	Ad5 Vx-SIV versus All others	-0.15	0.65	(-1.43, 1.12)	-0.15	0.63	(-1.38, 1.08)	0.58	0.39	(-0.19, 1.34)
Heterogeneous	Ad5 Vx-SIV versus Ad5 Vx-empty	-3.60	5.44	(-14.26, 7.06)	-3.50	5.22	(-13.73, 6.73)	0.65	0.41	(-0.15, 1.44)
	Ad5 Vx-SIV + Vx-SIV versus All others	-0.47	0.77	(-1.98, 1.05)	-0.45	0.74	(-1.89, 0.99)	0.28	0.46	(-0.61, 1.18)
	Ad5 Vx-SIV versus All others	-0.16	0.76	(-1.64, 1.33)	-0.15	0.73	(-1.59, 1.28)	0.58	0.41	(-0.23, 1.38)

Akaike information criterion with a correction for finite sample sizes (AIC_C) of the homogeneous and the heterogeneous infection risk models including the main effects for the five vaccine groups and dose levels and their interactions for the highest dose levels are 212.12 and 216.03, respectively. $AIC_C = 2k - 2\hat{\ell} + 2k(k+1)/(n-k-1)$, where k is the number of parameters, n is the number of animals, and $\hat{\ell}$ is the estimated log likelihood. Smaller value is more preferred.

The heterogeneous infection risk model developed in this article is computationally advantageous. It leads to a simple closed form for the marginal likelihood function. This allows us to avoid numerical integrations that are oftentimes required for other types of random effects models. Moreover, we are able to derive vaccine efficacy as a simple function of regression parameters.

Further research in the statistical methods is warranted to accommodate more complex but realistic scenarios arising in RLD experiments. One of the worthwhile avenues of research is to relax the assumption of no “memory” of challenge history (Regoes and others, 2005; Regoes, 2012). As challenges proceed, the differences in vaccine efficacy between vaccinated and unvaccinated groups will be diminished if previous challenges can cause immunization. Time-varying markers reflecting the potential immunization status of each animal might help investigating this issue. Identifying and modeling unintended immunization resulting from past challenges is a difficult task that requires future development in both scientific understanding and statistical techniques.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at <http://biostatistics.oxfordjournals.org>.

ACKNOWLEDGMENTS

Conflict of Interest: None declared.

FUNDING

This work was supported by the National Institutes of Health (P30 CA015704, R01 GM106177, R37AI05465, and R01 CA152089).

REFERENCES

- CONAWAY, M. R. (1990). A random effects model for binary data. *Biometrics* **46**(2), 317–328.
- COULL, B. A., HOUSEMAN, E. A. AND BETENSKY, R. A. (2006). A computationally tractable multivariate random effects model for clustered binary data. *Biometrika* **93**(3), 587–599.
- ELLENBERGER, D., OTTEN, R. A., LI, B., AIDOO, M., RODRIGUEZ, I. V., SARIOL, C. A., MARTINEZ, M., MONSOUR, M., WYATT, L., HUDGENS, M. G. and others. (2006). HIV-1 DNA/MVA vaccination reduces the per exposure probability of infection during repeated mucosal SHIV challenges. *Virology* **352**(1), 216–225.
- GARCÍA-LERMA, J. G., OTTEN, R. A., QARI, S. H., JACKSON, E., CONG, M., MASCOTRA, S., LUO, W., KIM, C., ADAMS, D. R., MONSOUR, M. and others. (2008). Prevention of rectal SHIV transmission in macaques by daily or intermittent prophylaxis with emtricitabine and tenofovir. *PLOS Medicine* **5**(2), e28.
- GOLDMAN, N. AND WHELAN, S. (2000). Statistical tests of γ -distributed rate heterogeneity in models of sequence evolution in phylogenetics. *Molecular Biology and Evolution* **17**(6), 975–978.
- HUDGENS, M. G. AND GILBERT, P. B. (2009). Assessing vaccine effects in repeated low-dose challenge experiments. *Biometrics* **65**(4), 1223–1232.
- HUDGENS, M. G., GILBERT, P. B., MASCOLA, J. R., WU, C. D., BAROUCH, D. H. AND SELF, S. G. (2009). Power to detect the effects of HIV vaccination in repeated low-dose challenge experiments. *Journal of Infectious Diseases* **200**(4), 609–613.
- KALBFLEISCH, J. D. AND PRENTICE, R. L. (2011). *The Statistical Analysis of Failure Time Data*, Volume 360. Hoboken, New Jersey: John Wiley & Sons.

- MOERBEEK, M. (2012). Sample size issues for cluster randomized trials with discrete-time survival endpoints. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences* **8**(4), 146–158.
- PAN, Z. AND LIN, D. Y. (2005). Goodness-of-fit methods for generalized linear mixed models. *Biometrics* **61**(4), 1000–1009.
- QURESHI, H., MA, Z. M., HUANG, Y., HODGE, G., THOMAS, M. A., DIPASQUALE, J., DESILVA, V., FRITTS, L., BETT, A. J., CASIMIRO, D. R. *and others*. (2012). Low-dose penile SIVmac251 exposure of rhesus macaques infected with adenovirus type 5 (Ad5) and then immunized with a replication-defective Ad5-based SIV Gag/Pol/Nef vaccine recapitulates the results of the phase IIb Step trial of a similar HIV-1 vaccine. *Journal of Virology* **86**(4), 2239–2250.
- R CORE TEAM. (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- REGOES, R. R. (2012). The role of exposure history on HIV acquisition: insights from repeated low-dose challenge studies. *PLOS Computational Biology* **8**(11), e1002767.
- REGOES, R. R., LONGINI, I. M., FEINBERG, M. B. AND STAPRANS, S. I. (2005). Preclinical assessment of HIV vaccines and microbicides by repeated low-dose virus challenges. *PLOS Medicine* **2**(8), e249.
- REYNOLDS, M. R., WEILER, A. M., PIASKOWSKI, S. M., KOLAR, H. L., HESSELL, A. J., WEIKER, M., WEISGRAU, K. L., LEÓN, E. J., ROGERS, W. E., MAKOWSKY, R. *and others*. (2010). Macaques vaccinated with simian immunodeficiency virus SIVmac239 δ nef delay acquisition and control replication after repeated low-dose heterologous SIV challenge. *Journal of Virology* **84**(18), 9190–9199.
- ROBB, M. L., RERKS-NGARM, S., NITAYAPHAN, S., PITISUTTITHUM, P., KAEWKUNGWAL, J., KUNASOL, P., KHAMBOONRUANG, C., THONGCHAROEN, P., MORGAN, P., BENENSON, M. *and others*. (2012). Ad hoc analysis of behavior and time as co-variates of the Thai phase III efficacy trial: RV 144. *The Lancet Infectious Diseases* **12**(7), 531.
- RODRIGUEZ, G. AND ELO, I. (2003). Intra-class correlation in random-effects models for binary data. *The Stata Journal* **3**(1), 32–46.
- SCHEIKE, T. H. AND JENSEN, T. K. (1997). A discrete survival model with random effects: an application to time to pregnancy. *Biometrics* **53**(1), 318–329.
- SELF, S. G. AND LIANG, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* **82**(398), 605–610.
- ZHANG, S. AND AHN, C. (2011). Adding subjects or adding measurements in repeated measurement studies under financial constraints. *Statistics in Biopharmaceutical Research* **3**(1), 54–64.

[Received December 10, 2013; revised July 7, 2014; accepted for publication July 9, 2014]