Biostatistics (2022) **23**, 2, *pp.* 507–521
doi: 10.1093/biostatistics/kxaa037
Advance Access publication on September 24, 2020

Immune correlates analysis using vaccinees from test negative designs

DEAN A. FOLLMANN*

Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, Bethesda MD

dean.follmann@nih.gov

LORI DODD

Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, Bethesda MD

Summary

Determining the effect of vaccine-induced immune response on disease risk is an important goal of vaccinology. Typically, immune correlates analyses are conducted prospectively with immune response measured shortly after vaccination and subsequent disease status regressed on immune response. In outbreaks and rare disease settings, collecting samples from all vaccinees is not feasible. The test negative design is a retrospective design used to measure vaccine efficacy where symptomatic individuals who present at a clinic are assessed for relevant disease (cases) or some other disease (controls) and vaccination status ascertained. This article proposes that test negative vaccinees have immune response to vaccine assessed both for relevant (e.g., Ebola) and irrelevant (e.g., vector) proteins. If the latter immune response is unaffected by active (Ebola) infection, and is correlated with the relevant immune response, it can serve as a proxy for the immune response of interest proximal to infection. We show that logistic regression using imputed immune response as the covariate and case disease as outcome can estimate the prospective immune response slope and detail the assumptions needed for unbiased inference. The method is evaluated by simulation under various scenarios including constant and decaying immune response. A simulated dataset motivated by ring vaccination for an ongoing Ebola outbreak is analyzed.

Keywords: Imputation; Likelihood, Logisitic regression; Regression calibration.

1. Introduction

The test negative design is a kind of case–control study used to estimate vaccine efficacy (Fukushima and Hirota, 2017). Originally proposed for pneumococcal vaccination Broome *and others* (1980), it has been used extensively for influenza vaccine see Jackson and Nelson (2013) as well as rotavirus vaccine (Boom *and others*, 2010). The basic design uses symptomatic subjects who come to a clinic with symptoms that could be either the vaccine preventable case disease (say Ebola) or a different control disease (non-Ebola). Vaccination status is ascertained and vaccine efficacy (VE) calculated as one minus the odds of vaccine among cases divided by the odds of vaccine among controls. This equals the prospective vaccine

*To whom correspondence should be addressed.

Published by Oxford University Press 2020. This work is written by US Government employees and is in the public domain in the US.

efficacy due to the invariance of the odds ratio for prospective or case–control sampling. The test negative study requires that all symptomatic vaccinees have the same probability of a clinic visit, an analogous condition for the symptomatic unvaccinated subjects, and that exposure be the same for all vaccinated and unvaccinated subjects. Numerous authors have investigated the sensitivity of the test negative design to these assumptions Sullivan *and others* (2016), Westreich and Hudgens (2016), Fukushima and Hirota (2017), and Lewnard *and others* (2018).

An important goal in vaccine development is to evaluate how vaccine-induced immune response correlates with risk of disease, known as a "correlate of risk" (Qin *and others*, 2007). Identification of a specific immune response that is associated with low disease risk provides a target for vaccine development. A common approach to evaluate immune correlates is to measure immune response to the vaccine shortly after vaccination determine who gets infected during prospective follow-up and then use logistic regression to predict the probability of disease as a function of measured immune response. Ideally, such an approach would be used in outbreak settings, see Halloran *and others* (2020), but at times it may not be feasible to prospectively collect samples. In addition, for rare diseases in non-outbreak settings, an enormous number of samples may need to be collected for traditional correlates analysis. In other settings, the durability of a vaccine is of interest and thus the immune response at the time of exposure is of interest though this may be many years after vaccination.

This article details an approach to immune correlates analysis under a test negative design using samples collected from vaccinees proximal to exposure. To fix ideas, consider the rVSV vaccine for Ebola which induces an immune response to Ebola antigens say X as well as to the Vesicular Stomatitis Vaccine (VSV) vector, say W. Naively, one might measure X among the symptomatic vaccinees who visit a clinic and fit a logistic regression with outcome being Ebola and covariate X. The problem is that Ebola immune response is likely massive in those with Ebola infection. What we really want is the pure vaccine-induced X just prior to exposure in both Ebola cases and controls. As a proxy, we can use W at the clinic visit which should be similar to W at exposure. Provided X and W are correlated, we can use W to impute the vaccine-induced X at exposure. We can then fit a logistic regression with Ebola as the outcome using the expectation of X conditional on W, say $\widehat{X}(W)$, as the covariate; also known as regression calibration. If the time interval from vaccination to exposure is relatively long, the substantial decay of the immune response over time may need to be addressed. This can be accomplished either by stratification, or flexible parametric modeling of risk over study time, provided the disease is rare. We evaluate the performance of the method via simulation and illustrate its use by a simulated dataset meant to reflect the recent Ebola outbreak in the Democratic Republic of the Congo where the rVSV Ebola vaccine was deployed using ring vaccination.

2. FORMULATION AND MODELS

2.1. *Constant immune response*

An important goal in vaccine development is to assess how immune response to the vaccine, X, is associated with the risk of infection/disease Y. In a randomized trial, vaccinees ($Z = 1$) have an immune response X measured shortly after vaccination. For now, assume that X does not change. Volunteers are followed for a period of time $[0, \tau]$, and disease status is recorded. With a test negative design, we need both the case disease and a different (control) disease for reference while the undiseased are not used. We thus define the disease of interest (case), a different disease (control), or undiseased which we denote by $Y = 1, 0, 2$, respectively. We need three possibilities for later use in a test negative design. If both diseases are acquired, we use the case disease. We assume that among vaccinees

$$P(Y = 1|X, Z = 1) = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)}, \tag{2.1}$$

and that

$$P(Y=0|X,Z=1)=\{1-P(Y=1|X,Z=1)\}\pi=\frac{1}{1+\exp(\beta_0+\beta_1 X)}\pi, \tag{2.2}$$

where π is the conditional probability that a vaccinee who avoids the case disease develops the control disease. Note that π is assumed free of X as it should not have an effect on acquiring the control disease. An immune correlates analysis relating risk of disease to immune response can be conducted by fitting (2.1) in vaccinees using as outcome the indicator that $Y=1$ (versus $Y\neq 1$). For developmental simplicity, we do not incorporate baseline covariates in (2.1) though one can.

For certain settings, X may not be measured shortly after vaccination. For rare diseases it may be prohibitively costly, or in outbreak settings it may not be logistically feasible. In other settings, the durability of a vaccine is unknown and thus X at exposure is of interest even though it may be many years since vaccination. Or in outbreak settings, it may not be logistically feasible to draw samples. In all these situations, a correlates analysis using the immune response proximal to exposure would be of great interest.

A design that collects information proximal to exposure is the test negative which takes symptomatic patients ($Y=0$ or $Y=1$) who present at a clinic and tests them for the disease of interest and records vaccination status. Those testing positive are classified as cases ($D=1$) while those testing negative are classified as controls ($D=0$). Formally, $D=1$ if $Y=1$ and $A=1$, while $D=0$ if $Y=0$, and $A=1$, where A is the indicator a person with either disease (i.e., $Y=0$ or 1) arrives at the clinic. Overall vaccine efficacy is estimated as

$$1-\frac{\widehat{P}(Z=1|D=1)/\{1-\widehat{P}(Z=1|D=1)\}}{\widehat{P}(Z=1|D=0)/\{1-\widehat{P}(Z=1|D=0)\}}=1-\frac{\widehat{P}(D=1|Z=1)/\{1-\widehat{P}(D=1|Z=1)\}}{\widehat{P}(D=1|Z=0)/\{1-\widehat{P}(D=1|Z=0)\}}, \tag{2.3}$$

where Z is the vaccine indicator see Guolo (2008). The second equality follows from the invariance of the odds to prospective and retrospective sampling. The test negative design can be much more efficient than a prospective study, but does require more assumptions.

Suppose, we somehow knew X in all vaccinees who arrived at the clinic. Let $P(A=1|Y=y)=\theta_y$ be the arrival probability for an symptomatic vaccinee with disease status $y=0,1$. Based on (2.1) and (2.2), we can derive the probability that a symptomatic vaccinee who arrives at the clinic has the case disease:

$$\begin{aligned}
&P(D=1|X,D=1\cup D=0,Z=1) && (2.4)\\
&=\frac{P(D=1|X,Z=1)}{P(D=1|X,Z=1)+P(D=0|X,Z=1)}\\
&=\theta_1\frac{\exp(\beta_0+\beta_1X)}{1+\exp(\beta_0+\beta_1X)}\left(\theta_1\frac{\exp(\beta_0+\beta_1X)}{1+\exp(\beta_0+\beta_1X)}+\theta_0\pi\frac{1}{1+\exp(\beta_0+\beta_1X)}\right)^{-1} && (2.5)\\
&=\frac{\theta_1\exp(\beta_0+\beta_1X)}{\theta_1\exp(\beta_0+\beta_1X)+\theta_0\pi}\\
&=\frac{\exp(\omega+\beta_1X)}{1+\exp(\omega+\beta_1X)}, && (2.6)
\end{aligned}$$

where $\omega=\beta_0+\log\{\theta_1/(\pi\theta_0)\}$. Thus, if we fit a simple logistic regression model with X as covariate among vaccinees in the clinic, we can recover the slope β_1 from (2.1). Note that this obtains even if we allow cases and controls to have different arrival probabilities. The usual test negative design requires

$\theta_0 = \theta_1$ for all vaccinees. We next consider the actual setting where X must be derived from measurements at arrival. Figuring out how to get X from test negative data is the major contribution of this article.

In practice, we can measure X when subjects arrive at the clinic, say $X(t_A)$ where t_A is the arrival time. Naively fitting (2.6) using the clinic $X(t_A)$ is problematic. For vaccinees who present with the control disease ($D = 0$), $X(t_A)$ should be close to $X(t_I)$, where t_I is just prior to infection as control disease should have little effect on the vaccine-induced immune response to the relevant vaccine antigen. But the vaccinated who present with the disease of interest ($D = 1$) are likely to have quite high $X(t_A)$ due to active infection. So $X(t_A)$ at arrival is different from $X(t_I)$ for vaccinees with $D = 1$.

However, certain vaccines (e.g., vector based vaccines) also induce an immune response to irrelevant antigens (e.g., the vector), say W, which should be relatively unaffected by active infection. To fix ideas, suppose that X, W achieve stable values shortly after vaccination and remain constant in $D = 0$ diseased controls. If X, W are correlated prior to exposure, then one could predict the unadulterated X at exposure using the W at presentation. One could then use $\widehat{X}(W)$, the imputed immune response to relevant vaccine antigens at exposure, in lieu of X. This works if W is unaffected by case or control disease and X, W are unaffected by control disease. But these requirements can be weakened. Suppose that for both groups, we observe $W' = W + E_W$ and for those with control disease we observe $X' = X + E_X$ where E_W, E_X are errors (which can be correlated and have non-zero means). Thus W, X are the pre-infection values while W', X' are the values observed at the clinic when subjects are diseased. If E_W differs for case and control disease, however, the control model fitted on W', X' does not apply to cases and a more complex approach would be required.

This strategy is displayed in Figure 1. The diamonds are actual $X(t)$ while the two solid bent lines are linear interpolations of the IgG antibody response to the outside of the Ebola virus i.e., Ebola glycoprotein (GP) following rVSV vaccination from two randomly selected subjects in Prevail 1, an immunogenicity trial of the rVSV vaccine Kennedy *and others* (2016). Immune response was measured at baseline, 1 week, 1 month, 6 months, and 1 year (diamonds). While IgG response to vector W was not measured, we illustrate interpolated hypothetical values by the dashed bent lines. We pretend that these two subjects were infected, with, respectively, malaria and Ebola. They became symptomatic and arrived at a clinic a few days after productive exposure. The W, X values for the malaria patient at arrival are similar to the values at exposure, but for the Ebola patient, only W is similar at exposure and arrival while X (dashed line) is massive from active Ebola infection., The bottom panel illustrates hypothetical data with a correlation of 0.70 between W, X that might have been obtained from a sample of vaccinated controls who arrived at the clinic. From this relationship, we can impute X at infection for the Ebola patient.

More formally, from the controls we can fit the model

$$X = \gamma_0 + \gamma_1 W + e, \tag{2.7}$$

where e is an error term. Using the estimated parameters, we impute X,

$$\widehat{X}(W) = \hat{\gamma}_0 + \hat{\gamma}_1 W$$

and fit the logistic regression model in the symptomatic vaccinees

$$P(D = 1|W, Z = 1) = \frac{\exp\{\omega_0 + \beta_1 \widehat{X}(W)\}}{1 + \exp\{\omega_0 + \beta_1 \widehat{X}(W)\}}. \tag{2.8}$$

This imputation is known as regression calibration and is a simple way to correct for measurement error in a covariate. Note that we impute X even in the vaccinated controls where X is known and thus treat both cases and controls in the same way. Using imputation in all those with $D = 1$ and X directly in all

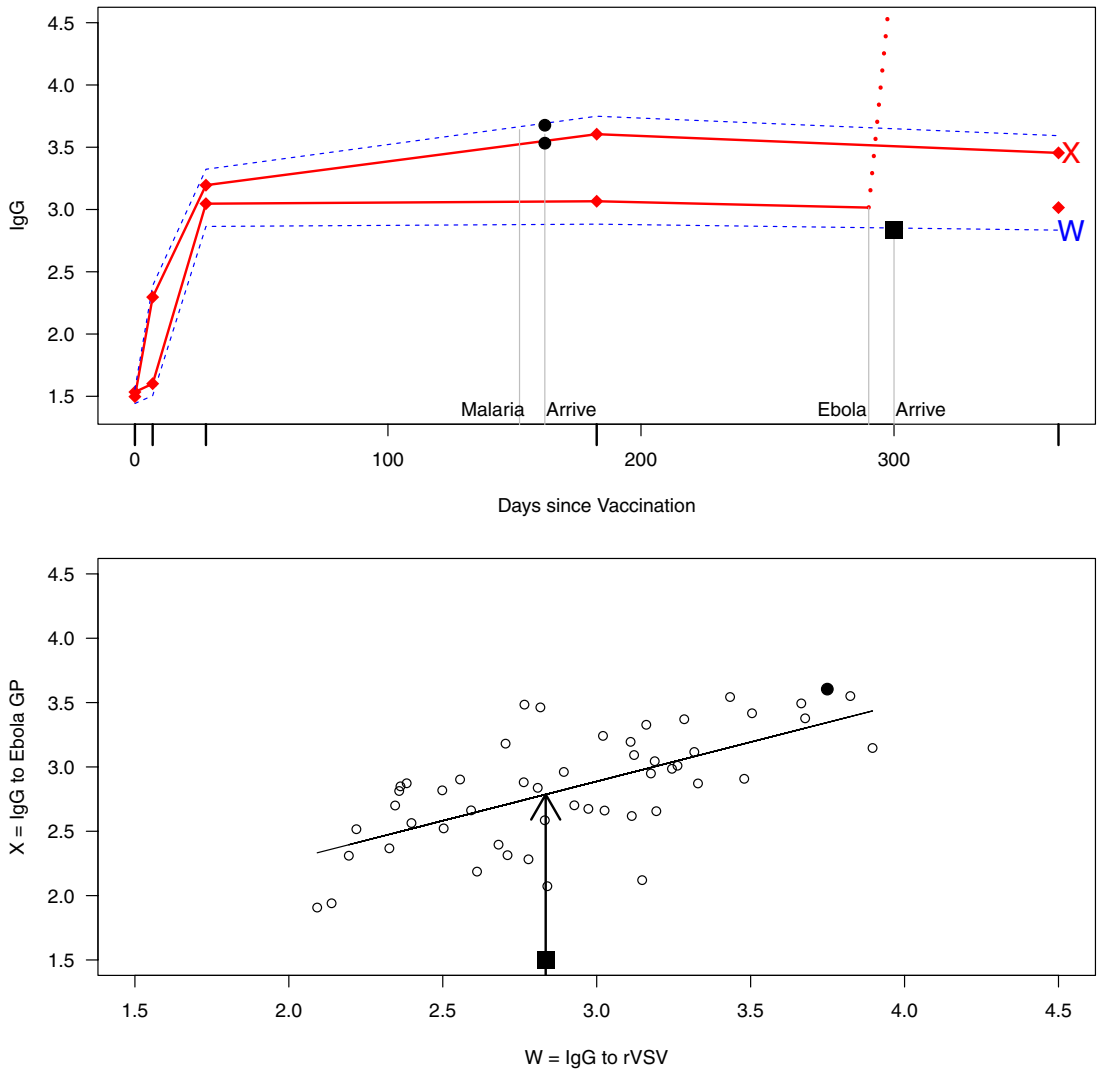

Fig. 1. Top panel: IgG immune responses over time for two vaccinated volunteers from Prevail 1. The solid bent line is the linear interpolation of X = IgG to Ebola GP measured at weeks 0, 1, 4, 26, and 52 (red diamonds). The dotted red line is hypothetical X to reflect Ebola infection. The dashed bent line is hypothetical W=IgG to rVSV vector. The subject infected with malaria arrives with $W(t_A), X(t_A)$ similar to those at infection. Only $W(t_A)$ is similar to the value prior to infection for the Ebola infected volunteer. Bottom panel: a scatter plot of $W(t_A), X(t_A)$ for the controls (circles) with imputation of $X(t_I)$ for the Ebola case (solid square).

those with $D = 0$ leads to substantial bias. Because we use $\widehat{X}(W)$, which is estimated instead of X, which is fixed, standard errors from standard software fitting (2.8) are likely too small. A simple remedy is to use the bootstrap or use derived standard errors, see Rosner *and others* (1990).

Instead of using an imputed X for regression calibration, one can perform a correlates analysis using W directly. While W is not part of the mechanism of protection, "correlates" analyses are often done with immune responses that may not be causal but are presumably related to the causal mechanism. As an example, smallpox vaccination involves scraping the arm with antigen. A vaccine "take" is recorded if a

pox forms. While the scar itself does not protect, it is a proxy for a robust relevant immune response to the vaccine and risk of disease by scar or "take" is of interest even though it is clearly a non-mechanistic correlate of risk using the nomenclature of Plotkin and Gilbert (2012). Use of W directly is simple, avoids the measurement error issue, requires no imputation, and has slightly better power than use of imputed $\widehat{X}(W)$, as we will see. Use of W by itself, however, is not easily interpretable unless as a proxy for vaccine "take," i.e., $X = W$ is binary and $W = X = 0$ is like being unvaccinated.

2.2. *Waning immune response*

The above development is appropriate if the immune responses X, W are stable during the followup period for the cohort study. In general, antibody decay over time and the above approach can lead to bias. To see this, suppose that nearly everyone was vaccinated prior to an outbreak which exploded and then waned so the risk of the case disease decreased over time, coinciding with the antibody decay. Further suppose that the control disease rate was constant over time. Then even if X was unrelated to Y we would associate low X with low risk and we would tend to estimate a positive β_1. And the problem is even more complicated as, in general, people will be vaccinated at various times and the case disease rate might vary with time.

Let $t \in [0, T]$ be the time since the start of the test negative study. To develop the time varying $X(t)$ setting more formally, we will assume that the hazard for a case disease arriving at time t with covariate $X(t)$ is given by

$$\lambda_1(t) = \lambda_{10}(t) \exp\{X(t)\beta_1\}\theta_1(t), \tag{2.9}$$

where $\lambda_{10}(t) \exp\{X(t)\beta_1\}$ is the hazard for case disease a little before t and $\theta_1(t)$ is the probability of arriving at the clinic at time t, given a case infection just prior to t.

We analogously assume that the hazard for the control disease is independent of $X(t)$ and arbitrary, as is the instantaneous probability of arriving. Thus, the hazard for a person with control disease arriving at the clinic is

$$\lambda_0(t) = \lambda_{00}(t)\theta_0(t).$$

Recall that the probability of an event in a small interval [t,t+Δ), given no event prior to t, is approximately $\lambda(t)\Delta$ for an arbitrary hazard function $\lambda()$. Under a "rare" disease assumption, we can calculate the probability that a given vaccinee who showed up at time t was a case as

$$\begin{aligned} P\{D(t) = 1|X(t), D(t) = 1 \cup D(t) = 0, Z = 1\} &\approx \frac{\lambda_{10}(t) \exp\{X(t)\beta_1\}\theta_1(t)\Delta}{\lambda_{00}(t)\theta_0(t)\Delta + \lambda_{10}(t) \exp\{X(t)\beta_1\}\theta_1(t)\Delta} \\ &= \frac{\exp\{\omega_{0t} + X(t)\beta_1\}}{1 + \exp\{\omega_{0t} + X(t)\beta_1\}}, \end{aligned} \tag{2.10}$$

where $D(t)$ is the indicator of case disease for an arrival at time t and

$$\omega_{0t} = \log\left(\frac{\lambda_{10}(t)\theta_1(t)}{\lambda_{00}(t)\theta_0(t)}\right).$$

Now ω_{0t} could be constant. One way this can happen is if $\lambda_{10}(t) = c\lambda_0(t)$ and $\theta_1(t) = a\theta_0(t)$ so $\omega_{0t} = \log(ca)$. If so, we can simply fit logistic regression to the n data points $\{D_i(t_i), X_i(t_i)\}, i = 1, \ldots, n$. In general, ω_{0t} will not be constant so that

$$\text{logit}\{P(D(t) = 1)\} = \omega_{0t} + \beta_1 X(t). \tag{2.11}$$

We might specify $\omega_{0t} = \omega_0 + \omega_1 I(t < T_M)$, where T_M is say the median followup time. Or we could fit a quadratic function of log(t). In practice, different approaches to specify ω_{0t} could be tried and the best one selected. Note that we only need know (or impute) $X(t)$ at the time of arrival to the clinic for each vaccinee.

3. Assumptions

The above approaches recover the prospective slope for X or W if the model assumptions are met. In this section, we delineate the assumptions when X is known for all, constant after time τ, the logistic-type model $P(Y = y|X, Z = 1)$ is correctly specified, and there is independence between exposure and X. The major issue is thus whether vaccinated cases (controls) arrive with fixed probability $\theta_1 = P(A = 1|Y = 1, Z = 1)$ ($\theta_0 = P(A = 1|Y = 0, Z = 1)$). Or equivalently, whether those who arrive are a representative sample of cases and controls respectively. We allow $\theta_0 \neq \theta_1$.

1. *The vaccinees who arrive with the control disease provide a representative sample of the distribution of immune response among those vaccinated without the case disease.*

This follows if X has no impact on the probability of acquiring the control disease and no impact on the probability of arriving at the clinic so $P(A = 1|Y = 0, Z = 1) = \theta_0$. Note we can allow $P(A = 1|Y = 0, Z = 1)$ to depend on latent or measured disease stratus or baseline covariates as long as they independent of X.

2. *The vaccinees who arrive with the case disease provide a representative sample of the distribution of immune response among those vaccinated with the case disease.*

$$\begin{aligned} P(A = 1|X, Y = 1, Z = 1) &= P(A = 1|S = 0, X, Y = 1, Z = 1)P(S = 0|X, Y = 1, Z = 1) + \\ &\quad P(A = 1|S = 1, X, Y = 1, Z = 1)P(S = 1|X, Y = 1, Z = 1) \\ &= P(A = 1|S = 0, Y = 1, Z = 1)P(S = 0|X, Y = 1, Z = 1) + \\ &\quad P(A = 1|S = 1, Y = 1, Z = 1)P(S = 1|X, Y = 1, Z = 1). \end{aligned} \tag{3.12}$$

The last equality follows if X has no additional impact on A given severity which seems a reasonable assumption. Using the last equality, we can see that $P(A = 1|X, Y = 1, Z = 1)$ does not depend on X if either $P(A = 1|S, Y = 1, Z = 1)$ is free of S or $P(S|X, Y = 1, Z = 1)$ is free of X. We discuss each in turn.

Now $P(A = 1|S, Y = 1, Z = 1) = P(A = 1|Y = 1, Z = 1)$ if severity does not impact health seeking behavior. This might occur if all cases are equally encouraged to arrive at a clinic and all cases are equally compliant, or if the severity gradient is modest enough that it does not change behavior.

Next consider when $P(S|X, Y = 1, Z = 1) = P(S|Y = 1, Z = 1)$, i.e., X has no impact on severity given a vaccinee is infected. While this obtains if the vaccine's mechanism of action is all-or-none, it is a weaker condition. An all-or-none vaccine implies that the distribution of severity is unchanged by vaccination so that $F(S|Y = 1, Z = 1) = F(S|Y = 1, Z = 0)$, where Z is the vaccine indicator. But we only require $F(S|Y = 1, X, Z = 1) = F(S|Y = 1, Z = 1)$ which allows $F(S|Y = 1, Z = 1) \neq F(S|Y = 1, Z = 0)$.

If neither assumption is plausible but case disease severity is observed, there is a remedy provided disease follows the prospective logistic regression model:

$$P(Y = 1, S = s|X, Z = 1) = \frac{\exp(\beta_{0s} + \beta_1 X)}{1 + \exp(\beta_{0s} + \beta_1 X)}, \tag{3.13}$$

for $s = 0, 1$. Then even if X affects the arrival probability through the observable disease severity, recovery of β_1 using data D, S, X follows from arguments analogous to those used to derive (2.6). One can show that

(3.13) implies the distribution of disease severity among the vaccinated infecteds, $P(S = s|Y = 1, X)$, varies with X reflecting a kind of disease modifying vaccine.

These conditions can be relaxed using baseline covariates V. If the prospective model (2.1) controls for confounding by specifying $\text{logit}\{P(Y = 1|X, V, Z = 1)\} = \beta_0 + \beta_1 X + \beta_2 V$, then the above arguments go through, even if the arrival probabilities satisfy $\theta_z(V) = \nu_z \exp(\gamma_Z V)$. One can show that $\text{logit}\{P(D = 1|X, V, D = 1 \cup D = 0, Z = 1)\} = \omega(V) + \beta_1 X$ with $\omega(V) = \beta_0 + (\beta_2 + \gamma_1 - \gamma_0)V + \log\{\nu_1/(\pi\nu_0)\}$, and thus β_1 remains the coeffience for X.

To summarize, assuming the verity of (3.12) and a correctly specified model for $P(Y = 1|X, Z = 1)$, we can recover β_1 even if θ_0, θ_1 depend on V unless disease severity is unobservable, $P(S|Y = 1, X, Z = 1)$ depends on X, and $P(A = 1|S, Y = 1, Z = 1)$ depends on S.

4. SIMULATIONS

4.1. *Constant immune response*

We illustrate performance for the simple case where antibody does not decay. We prospectively generate data for vaccinees which are then sampled retrospectively as in a test negative design. We assume a population of size N vaccinees and generate the case disease according to

$$P(Y = 1|X, Z = 1) = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)},$$

and sequentially generate control disease among the case uninfecteds as

$$P(Y = 0|X, Z = 1) = \left(1 - \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)}\right)\pi.$$

We specify X as Gaussian with mean= 3.00 and standard deviation =0.50 as in the Prevail 1 vaccine trial Kennedy *and others* (2016). We generate

$$W = 3 + \rho(X - 3) + \sigma(1 - \rho^2)U,$$

where U is standard normal. Thus, X, W are bivariate normal with common mean 3.00, common standard deviation 0.50, and correlation ρ. We set $\rho = 0.40$, 0.70, and 1.00. For reference, a study estimated a correlation of 0.70 between ELISA OD immune response readouts for VSV proteins (W) and Ebola proteins (X) at 56 days post vaccination Poetsch *and others* (2018).

We specify (β_0, β_1) as $(-5.51, 0.00)$, $(-4.05, -0.60)$, and $(-1.12, -1.81)$. These correspond to no, modest, and strong effects of X. To help interpret β_1, the ratios of risk of case disease for the 1st versus 8th octile of X are 1, 2, and 10, respectively, for $\beta_1 = 0.00, -0.60, -1.81$. We set the conditional risk of control disease as $\pi = 0.01$, $N = 150\,000$, and $\theta_0 = \theta_1 = 0.50$ so that vaccinated cases and controls arrive with equal probability.

We evaluate estimation of the model

$$P(D = 1|\text{IR}, Z = 1) = \frac{\exp(\omega_0 + \beta_1 \text{IR})}{1 + \exp(\omega_0 + \beta_1 \text{IR})}, \tag{4.14}$$

where IR, the immune response, is X, $\widehat{E}(X|W)$, or W

Table 1 presents the results using 10 000 simulated studies per scenario. Columns 4 and 5 provide the sample mean and variance of $\hat{\beta}_1$. Column 6 reports the rejection rate for the two-sided $\alpha = 0.05$ Wald test

Table 1. *Simulated performance of logistic regression using different covariates for IR immune response; $X, \widehat{X}(W)$, or W. Sample statistics for different parameters estimates are presented. $\overline{n}$ is the average number of symptomatic vaccinees who arrive and $\overline{n}_1$ the average number of vaccinees who arrive with the case disease. 1000 test negative designs are simulated for each row. The last row has $\theta_0 = 0.25, \theta_1 = 0.50$ all other rows have $\theta_0 = \theta_1 = 0.50$. Each row is based on 10 000 simulated studies.*

β_1	ρ	IR	$\overline{\widehat{\beta}_1}$	$S^2(\widehat{\beta}_1)$	% Reject	$\overline{n}$	$\overline{n}_1$
0.000	1.000	X	−0.002	0.019	0.053	1049	302
−0.600	1.000	X	−0.602	0.024	0.973	972	225
−1.810	1.000	X	−1.821	0.040	1.000	908	160
0.000	0.700	$\widehat{X}(W)$	−0.001	0.040	0.048	1048	302
−0.600	0.700	$\widehat{X}(W)$	−0.605	0.049	0.786	973	225
−1.810	0.700	$\widehat{X}(W)$	−1.826	0.076	1.000	909	160
0.000	0.400	$\widehat{X}(W)$	0.003	0.122	0.049	1049	302
−0.600	0.400	$\widehat{X}(W)$	−0.603	0.152	0.332	973	225
−1.810	0.400	$\widehat{X}(W)$	−1.827	0.232	0.974	909	160
0.000	0.700	W	−0.001	0.020	0.048	1048	302
−0.600	0.700	W	−0.423	0.024	0.790	973	225
−1.810	0.700	W	−1.274	0.035	1.000	909	160
0.000	0.400	W	0.001	0.019	0.048	1049	302
−0.600	0.400	W	−0.239	0.023	0.344	973	225
−1.810	0.400	W	−0.725	0.032	0.984	909	160
−0.600	1.000	X	−0.604	0.031	0.938	599	225

using the Monte Carlo standard error over the 10 000 simulated studies. We see that we have power of around 100% to test the effect of IR when $\beta_1 = -1.81$ for all scenarios, while for $\beta_1 = -0.60$ we need a correlation of 0.70 to approach 80% power. The type I error rate seems consistent with the $\alpha = 0.05$ used for testing.

Regression calibration recovers the true β_1 even with $\rho = 0.40$. In the Appendix, we derive the correct (marginalized) probability of $D = 1$ given W under our assumptions. We show that this curve as a function of W (with slope β_1) is virtually identical to the regression calibration curve (with the same slope β_1), irrespective of ρ. This fidelity explains the excellent recovery of β_1 with regression calibration. Regression calibration does become more biased as $|\beta_1| \times \sigma_x$ increases, but for these simulations, $|\beta_1| \times \sigma_x$ is relatively small, Follmann *and others* (1999).

Intuitively, as ρ decreases estimation becomes more difficult. For $\beta_1 = -0.60$, the variance for $\hat{\beta}_1$ under regression calibration is 0.024, 0.049, and 0.152 for ρ = 1.00, 0.70, and 0.40, respectively. So the variance roughly doubles as ρ goes from 1.00 to 0.70 and triples as ρ goes from 0.70 to 0.40. In contrast, with direct use of W the variance of $\hat{\beta}_1$ stays constant at about 0.023 or 0.024 for all ρ with $\beta_1 = -0.60$. The use of W in lieu of X, of course, leads to a smaller estimate of β_1 but this is counterbalanced by the smaller variance. The upshot is that the power for testing $\beta_1 = 0$ is nearly identical for regression calibration and direct use of W for all scenarios.

In the Appendix, we provide additional simulations with sample sizes 1/10th and 10 times that of Table 1. The three approaches have less power and more variability with the smaller sample size and the opposite with larger sample size. Regression calibration continues to be nearly unbiased for the larger sample size, though it has some bias for the smaller sample size.

The above simulations were conducted under the standard assumptions of the test negative design. In the final row, we set $\theta_0 = 0.25$ while $\theta_1 = 0.50$ so that vaccinees with the control disease have half the arrival probability of the vaccinated cases, which violates the standard test negative assumption. We set $\beta_1 = -0.60$ and evaluate $IR = X$ so that the second line of Table 1, which has $\theta_0 = \theta_1 = 0.50$ provides a reference. As expected, we see the estimate of β_1 is still unbiased, though the variance for $\hat{\beta}_1$ is slightly increased due to fewer control vaccinees.

4.2. *Waning immune response*

In this subsection, we provide a brief evaluation of the model (2.11). To focus on the issue of antibody decay, we assume that $X(t) = W(t)$ is measured without error, and that everyone with either disease arrives at a clinic. We specify the instantaneous conditional risk for control disease using a Weibull hazard,

$$\lambda_0(t) = t^{\alpha_0 - 1} \exp(\beta_{00}),$$

where $t \in [0, T]$ is the time since start of the test negative study. When $\alpha_0 > 1$ (< 1) the hazard is increasing (decreasing). The instantaneous conditional risk of case disease is specified as

$$\lambda_1(t) = t^{\alpha_1 - 1} \exp\{\beta_{01} + \beta_1 X(t)\}.$$

For $X(t)$, we specify linear antibody decay according to a random effects model. For person i, the decay is

$$X_i(t) = b_{0i} + b_{1i} t + B_0 + B_1 t,$$

where b_{0i}, b_{1i} are independent normals with mean 0 and standard deviations σ_0, σ_1, respectively. The parameter σ_0 reflects both natural variation in immune response to vaccination as well as variation in vaccination times as those vaccinated long ago would tend to have smaller b_{0i}.

We generate times to case and control disease by the inverse cumulative distribution function method where we generate U, a uniform [0,1] random variable and then determine the disease arrival time t that solves $F_d(t) = U$ for $d = 0,1$. For the case disease, this is solved by numerically minimizing $\{F_1(t) - U\}^2$. For subjects who experience both case and control events, we use the first and discard the second. We set $N = 10\,000$ vaccinees and follow subjects until we observe 1000 first events.

For all scenarios, we set $\beta_{00} = -2.25$, $(\beta_{01}, \beta_1) = (2.18, -1.47)$, and $B_0 = 3.50$. Our reference case is $(\alpha_0, \alpha_1) = (1.00, 1.00)$, or constant risk of each disease and $\sigma_0 = 0.50$, $\sigma_1 = 0.10$. We then vary these parameters to see their impact on performance. The hazards and distribution of arrival times are graphed in the Supplementary material available at *Biostatistics* online.

We evaluate two different modeling strategies for (2.11). The first is to treat $\omega_{0t} = \omega_0$ as fixed with no dependence on time. The second is to evaluate five models and pick the one with the best Akaike Information Criterion (AIC). The five models are to treat ω_{0t} as constant (1), linear in t (2), quadratic in t (3), linear in $\log(t)$ (4), and quadratic in $\log(t)$. This second strategy is meant to mimic how a data analyst might address a non-constant ω_{0t}. To summarize performance, we calculate the sample mean and variance for the estimate of β_1 under modeling strategies 1 and 2. We also report the modal model choice of strategy 2 and the number of times that modal choice was chosen. Additionally, we report the number of cases and controls for the last data set.

Table 2 presents the summary results. We see that both modeling strategies 1 and 2 recover β_1 very well for scenario 1, our reference, and scenario 2, where $X(t)$ has more variability and a steeper mean decline. Both scenarios 1 and 2 have constant hazards for both case and control disease. For non-constant

Table 2. *Performance of logistic regression based strategies for dealing with the nuisance function ω_{0t}. Strategy 1: Fix ω_{0t} at β_0. Strategy 2: Choose ω_{0t} based on AIC. The sample average and variance over the simulations are given for $\hat{\beta}_1$ under the two strategies. AIC choice reports the modal choice for strategy 2 while n_1, n_0 are the number of cases and controls for the last simulated dataset. The true value of $\beta_1 = -1.47$. Each row is based on 1000 simulated studies.*

						Strategy 1		Strategy 2		AIC	# times		
Scenario	α_1	α_0	σ_0	σ_1	B_1	$\overline{\hat{\beta}_1}$	$S^2(\hat{\beta}1)$	$\overline{\hat{\beta}_1}$	$S^2(\hat{\beta}_1)$	Choice	Chosen	n_1	n_0
1	1.00	1.00	0.5	0.1	−0.5	−1.47	0.0204	−1.47	0.0209	1	735	426	574
2	1.00	1.00	1.0	0.5	−1.0	−1.48	0.0100	−1.48	0.0102	1	693	629	371
3	1.25	0.75	0.5	0.1	−0.5	−1.59	0.0268	−1.48	0.0271	4	738	235	765
4	0.75	1.25	0.5	0.1	−0.5	−1.35	0.0213	−1.48	0.0239	4	780	618	382
5	1.50	0.50	0.5	0.1	−0.5	−1.61	0.0858	−1.49	0.0927	4	770	66	934
6	0.50	1.50	0.5	0.1	−0.5	−1.29	0.0324	−1.48	0.0392	4	828	829	171

hazards, modeling strategy 1 is biased with bias of around ±10%. Modeling strategy 2 has minimal bias for non-constant hazards and picks ω_{0t} as a linear function of $\log(t)$ in about 70–80% of the simulations using the AIC criterion.

Strictly speaking, (2.10) obtains under a "rare" disease assumption. In the above simulations, 10% of the vaccinees acquired disease with no noticeable bias. We did further simulations and did observe about 7% bias for scenario 2 with 30% disease acquisition.

5. Example

In the 2018–2020, outbreak of Ebola in the Democratic Republic of the Congo a ring vaccination campaign was conducted. Local surveillance teams kept track of new Ebola cases which were then relayed to a WHO vaccination team that vaccinated the contacts and contacts of contacts of the index case. Several hundred thousand at risk people were thus vaccinated. Following vaccination, the vaccinees were monitored and encouraged to visit an Ebola transit/treatment center once any symptoms developed. The ring campaign used the rVSV vaccine which is estimated to have quite high efficacy though breakthrough infections occur, WHO (2019) and Henao-Restrepo *and others* (2015). The durability of the vaccine and the impact of immune response on risk of disease are unknown. Prospectively collecting samples in all vaccinees for correlates analysis has not been done.

To illustrate our approach, we simulate a dataset meant to loosely approximate a large Ebola ring vaccination campaign with a vaccine whose substantial efficacy is modulated by immune response. We assume that each person's immune response is relatively stable over time. We assume 150 000 at risk subjects are vaccinated and generate case and control infections as in the previous section with $(\beta_0, \beta_1) = (-1.12, -1.81)$. We set π, the conditional risk of control disease (e.g., malaria) equal to 0.005 and assume that vaccinated subjects arrive at the transit center with fixed probability $\theta_0 = \theta_1 = 0.50$, as one might expect from regular monitoring and encouragement. Note that with regular monitoring of all subjects in a ring, the assumption that $P(A|Y = 1, S, Z = 1) = P(A|Y = 1, Z = 1)$ for vaccinees may hold, thus even if rVSV does modify disease, the estimated risk of disease as a function of immune response should be unbiased.

To provide additional context, we assume 300 000 unvaccinated subjects are at risk (e.g., have first or second degree contact with an Ebola case) and generate case infections with probability 0.01. We generate control infections from this set with conditional probability 0.005 just as for the vaccinated. We assume

that the unvaccinated contacts arrive with probability 0.25 for both case and control disease, irrespective of severity. Finally, we assume that 20% of the arrivals occur within 35 days of contact with an Ebola case. For vaccinees, this should ensure a relatively stable vaccine-induced immune response at arrival.

For this single simulated dataset, a total of 114 unvaccinated contacts of Ebola cases arrived at least 35 days post contact. The proportion with Ebola Virus Disease (EVD) was 0.69. In contrast, 127 vaccinated contacts of an Ebola case arrived at least 35 days post contact with an EVD proportion of 0.28. Thus the overall test negative VE for these late arrivals is estimated as

$$1 - \frac{0.28/0.72}{0.69/0.31} = 0.82$$

The above VE estimate is unbiased under the standard test negative assumptions that vaccinees (unvaccinated) have fixed probability $P(A = 1|Z = 1)$ ($P(A = 1|Z = 0)$) of arriving at a transit center, irrespective of true disease status, disease severity, and equal exposure of vaccinees and unvaccinated. This latter assumption may be questioned in a ring study where the vaccinees are all known contacts of a case.

Figure 2 displays the data, jittered for the unvaccinated for whom we set X to 0, along with the fitted logistic regression model $\text{logit}\{\hat{P}(D = 1|\text{I}R, Z)\} = \widehat{\omega}_0 + \widehat{\beta}_1 IR$ where IR is the immune response covariate, X, $\widehat{X}(W)$, or W. For the unvaccinated, we draw a dashed line at the observed Ebola disease rate of 0.69 for reference. Table 3 provides the estimated parameters. We see that, as expected, using W results in a smaller $\hat{\beta}_1$ than use of $\widehat{X}(W)$ and that the Wald statistic for testing $\beta_1 = 0$ for W is identical to the Wald statistic based on $\widehat{X}(W)$, when we use the naive standard error. We also provide the bootstrap standard error and Wald statistic for regression calibration based on 1000 resamples. The resultant standard errors are slightly larger than the naive standard errors. The two-sided p-values for testing an effect of $\widehat{X}(W)$ and W on disease are, respectively, 0.02 and 0.01. Recall that the estimates of β_1 are unbiased even if the arrival probability for vaccinees depends on latent severity, provided the vaccine has an all-or-none mechanism. Estimates of β_1 are also unbiased even if the vaccine modifies latent disease severity, provided the arrival probability is the same for all vaccinees with the case disease.

To help interpret the magnitude of the effect of the imputed immune response on risk of disease, we can compare the ratio of the risk of disease at the 10th versus 90th percentile of the observed distribution of $\widehat{X}(W)$. These percentiles are 2.51 and 3.39, respectively, with associated probabilities of disease of 0.42 and 0.16 for these late vaccinated arrivals. Thus the risk is roughly 2.5 times greater at the 10th versus 90th percentile of $\widehat{X}(W)$ indicating a substantial effect of antibody on EVD risk.

6. DISCUSSION

This research arose from the design of an immune correlates study using test negative data for the 2018–2020 Democratic Republic of the Congo (DRC) Ebola outbreak. The main contribution of this article is the idea that an estimate of the effect of the relevant immune response X on the risk of disease can be estimated by using an irrelevant immune response W. We use the proxy W to impute X based on a prediction model that is estimated in vaccinated controls using samples collected upon arrival at a clinic. The method extends to settings where antibody substantially decays over time. Simulations demonstrated the feasibility of the proposed methods.

This approach applies beyond Ebola. The key requirement is that the vaccine of interest produce irrelevant immune responses that will be unaffected by the disease of interest, and that these immune responses are correlated with the immune response of interest. One such application is for extremely rare diseases such as the prevention of Zika-induced congenital birth defects. Suppose Puerto Rico begins Zika vaccination in 2025 and an outbreak sweeps the island in 2040. Since Zika-induced birth defects are

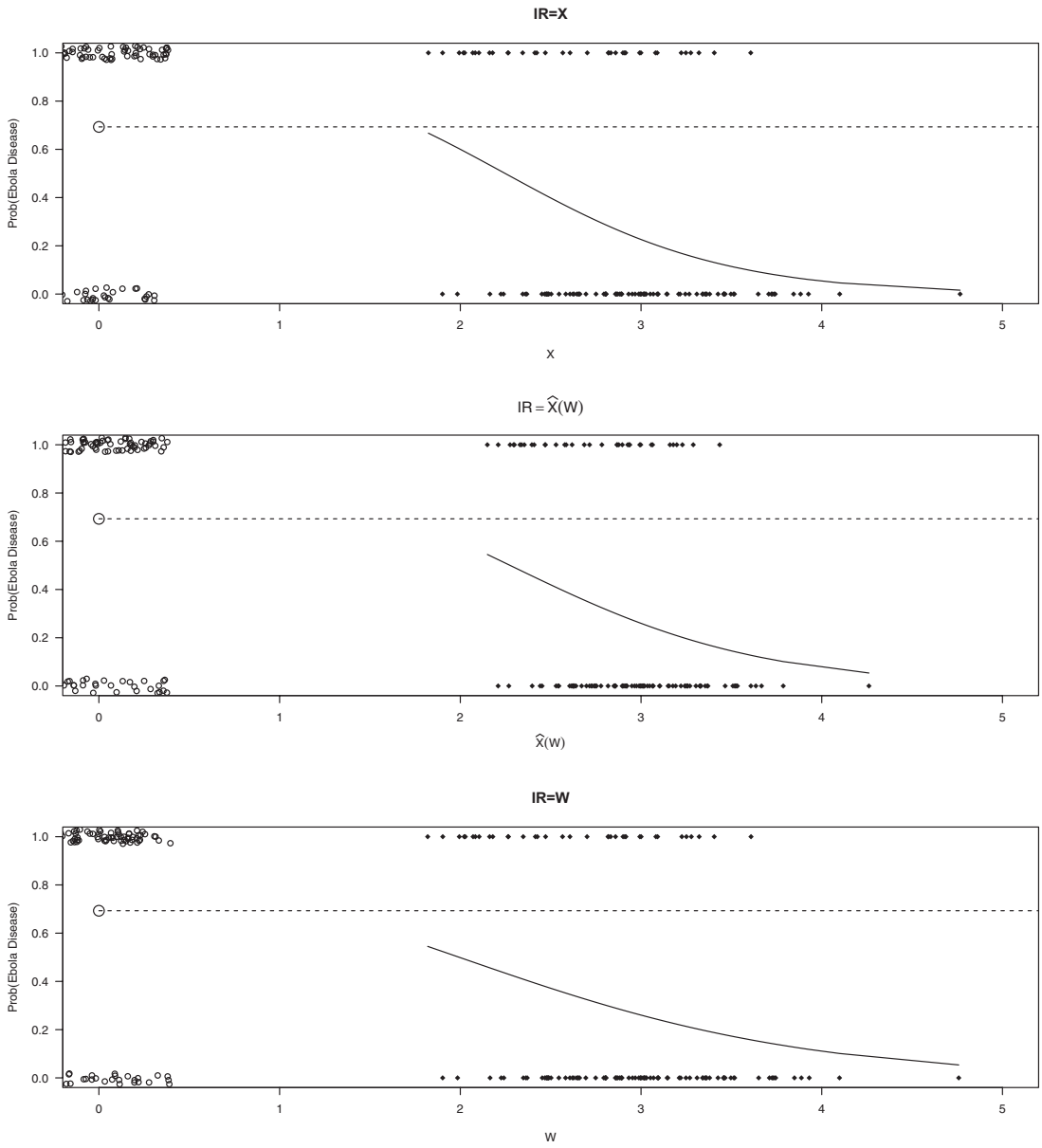

Fig. 2. Simulated data to illustrate an immune correlates analysis based on a test negative design for the 2018–2020 Ebola outbreak in the DRC. Unvaccinated subjects' data are given by jittered open circles near x-axis near 0 and whose Ebola disease rate of 0.69 is given by the dashed line. Vaccinated subjects' data are given by diamonds. The estimated probability of disease for vaccinees is given by the solid curve. The three panels correspond to using X, $\widehat{X}(W)$, and W as the immune response, respectively.

rare, a prospective immune correlates design would require long-term storage of samples from hundreds of thousands of women just after vaccination. Furthermore, the values collected in 2025 onward would not address the question of the immune response at Zika exposure in 2040. But using the methods of this article we could sample cases (mothers of children with birth defects who are antibody positive for

Table 3. *Parameter estimates for a simulated data set with 114 unvaccinated and 127 vaccinated subjects. We evaluate X, $\widehat{X}(W)$, and W as the immune response covariate.*

	IR = X			IR=$\widehat{X}(W)$					IR= W		
parm	est	se	Wald	est	se*	Wald*	se**	Wald**	est	se	wald
ω_0	3.68	1.28	2.86	3.28	1.68	1.95	1.70	1.93	2.07	1.20	1.72
β_1	−1.64	0.46	−3.54	−1.45	0.58	−2.49	0.60	−2.40	−1.04	0.42	−2.49

*Naive standard errors, **Bootstrap standard errors

non-vaccine Zika antigens) and controls (mothers of children with birth defects who are antibody negative for non-vaccine Zika antigens). We could then apply a test negative design to assess overall VE and also perform an immune correlates analysis with W and $\widehat{X}(W)$, providing that a W that predicts X were available.

We focused on $X(t_I)$ or the immune response proximal to infection. However, traditional immune correlates analyses focus on X shortly after vaccination. Our methods for time constant immune response directly apply as here $X = X(t_I)$. If antibody decay occurs, one could in principle apply (2.10) with $X(t)$ replaced by $\widehat{X}(t_v)$, the predicted value of X shortly after vaccination. This would require the development of models to predict $X(t_v)$.

The biggest challenge in this approach is whether the biology is amenable. Some vaccines won't induce a W at all. Other times, even if W is induced the correlation may be too weak for this method to work. This may be more of a problem with pre-existing immunity to either the antigens of the vaccine or the vector. On the other hand, it is not really required that W be an immune response to the vector. What is required is the existence of covariates $\mathbf{W}$ that predict X among vaccinees and that $\text{logit}\{P(Y = 1|X, \mathbf{W}, Z = 1)\} = \beta_0 + \beta_1 X$. Development of such methods is left to future work.

7. SOFTWARE

The code to reproduce the example is given at https://github.com/follmand/Test-Negative-Immune-Correlates.

SUPPLEMENTARY MATERIALS

Supplementary material is available at http://biostatistics.oxfordjournals.org.

ACKNOWLEDGMENTS

This work utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). We thank Jing Qin and Michael Fay for helpful comments.

Conflict of Interest: None declared.

REFERENCES

BOOM, J. A., TATE, J. E., SAHNI, L. C., RENCH, M. A., HULL, J. J., GENTSCH, J. R., PATEL, M. M., BAKER, C. J. AND PARASHAR, U. D. (2010). Effectiveness of pentavalent rotavirus vaccine in a large urban population in the United States. *Pediatrics* **125**, e199–e207.

BROOME, C. V., FACKLAM, R. R. AND FRASER, D. W. (1980). Pneumococcal disease after pneumococcal vaccination: an alternative method to estimate the efficacy of pneumococcal vaccine. *New England Journal of Medicine* **303**, 549–552.

FOLLMANN, D. A., HUNSBERGER, S. A. AND ALBERT, P. S. (1999). Repeated probit regression when covariates are measured with error. *Biometrics*, **55**, 403–409.

FUKUSHIMA, W. AND HIROTA, Y. (2017). Basic principles of test-negative design in evaluating influenza vaccine effectiveness. *Vaccine* **35**, 4796–4800.

GUOLO, A. (2008). A flexible approach to measurement error correction in case–control studies. *Biometrics* **64**, 1207–1214.

HALLORAN, M. E., LONGINI, I. M. AND GILBERT, P. B. (2020). Designing a study of correlates of risk for Ebola vaccination. *American Journal of Epidemiology* **189**, 747–754.

HENAO-RESTREPO, A. M., LONGINI, I. M., EGGER, M., DEAN, N. E., EDMUNDS, W. J., CAMACHO, A., CARROLL, M. W., DOUMBIA, M, DRAGUEZ, B., DURAFFOUR, S. *and others*. (2015). Efficacy and effectiveness of an RVSV-vectored vaccine expressing Ebola surface glycoprotein: interim results from the guinea ring vaccination cluster-randomised trial. *The Lancet* **386**, 857–866.

JACKSON, M. L. AND NELSON, J. C. (2013). The test-negative design for estimating influenza vaccine effectiveness. *Vaccine* **31**, 2165–2168.

KENNEDY, S. B., NEATON, J. D., LANE, H. C., KIEH, M. W. S., MASSAQUOI, M. B. F., TOUCHETTE, N. A., NASON, M. C., FOLLMANN, D. A., BOLEY, F. K., JOHNSON, M. P. *and others*. (2016). Implementation of an Ebola virus disease vaccine clinical trial during the Ebola epidemic in Liberia: design, procedures, and challenges. *Clinical Trials*, **13**, 49–56.

LEWNARD, J. A., TEDIJANTO, C., COWLING, B. J. AND LIPSITCH, M. (2018). Measurement of vaccine direct effects under the test-negative design. *American Journal of Epidemiology* **187**, 2686–2697.

PLOTKIN, S. A. AND GILBERT, P. B. (2012). Nomenclature for immune correlates of protection after vaccination. *Clinical Infectious Diseases* **54**, 1615–1617.

POETSCH, J. H., DAHLKE, C., ZINSER, M. E., KASONTA, R., LUNEMANN, S., RECHTIEN, A., LY, M. L., STUBBE, H. C., KRÄHLING, V., BIEDENKOPF, N. *and others*. (2018). Detectable vesicular stomatitis virus (VSV)-specific humoral and cellular immune responses following VSV-Ebola virus vaccination in humans. *The Journal of Infectious Diseases* **219**, 556–561.

QIN, L., GILBERT, P. B., COREY, L., MCELRATH, M. J. AND SELF, S. G. (2007). A framework for assessing immunological correlates of protection in vaccine trials. *The Journal of Infectious Diseases* **196**, 1304–1312.

ROSNER, B., SPIEGELMAN, D. AND WILLETT, W. C. (1990). Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *American Journal of Epidemiology* **132**, 734–745.

SULLIVAN, S. G., TCHETGEN TCHETGEN, E. J. AND COWLING, B. J. (2016). Theoretical basis of the test-negative study design for assessment of influenza vaccine effectiveness. *American Journal of Epidemiology* **184**, 345–353.

WESTREICH, D. AND HUDGENS, M. G. (2016). Invited commentary: beware the test-negative design. *American Journal of Epidemiology* **184**, 354–356.

WHO. (2019). *Preliminary Results on the Efficacy of rVSV-ZEBOV-GP Ebola Vaccine using the Ring Vaccination Strategy in the Control of an Ebola Outbreak in the Democratic Republic of the Congo: An Example of Integration of Research into Epidemic Response*. https://reliefweb.int/report/democratic-republic-congo/preliminary-results-efficacy-rvsv-zebov-gp-ebola-vaccine-using-ring.

[*Received March 10, 2020; revised August 10, 2020; accepted for publication August 15, 2020*]