The International Journal of Biostatistics

Volume 8, Issue 2

2012

Article 3

CAUSAL INFERENCE IN HEALTH RESEARCH

Bounds on the Effect of Vaccine Induced Immune Response on Outcome

Dean Follmann, *National Institute of Allergy and Infectious
Diseases*

Michael Fay, *National Institute of Allergy and Infectious
Diseases*

Recommended Citation:

Follmann, Dean and Fay, Michael (2012) "Bounds on the Effect of Vaccine Induced Immune Response on Outcome," *The International Journal of Biostatistics*: Vol. 8: Iss. 2, Article 3.

DOI: 10.2202/1557-4679.1348

Bounds on the Effect of Vaccine Induced Immune Response on Outcome

Dean Follmann and Michael Fay

Abstract

A major goal of vaccine development is the identification of immune responses that are responsible for vaccine efficacy. In theory, modest vaccines could be successfully improved by increasing such immune responses. And for a vaccine with a great benefit in one population, inducing such immune response in a different population could help one conclude the vaccine would have great benefit there. Such identification is tricky because the immune response to vaccination can only be measured in the vaccine group and thus immune responses might only be identifying individuals with a constitutional ability to remain uninfected, rather than being causal. Define the vaccine induced immune response as $X(1)$. The value $X(1)$ is a potential outcome; it is measured directly in vaccinees but unobserved in the placebo group. Our goal is to regress outcome on $X(1)$ separately in the vaccine and placebo groups and to see if the vaccine effect varies with $X(1)$. Put another way, our goal is to see if there is a vaccine by $X(1)$ interaction. Regression of outcome on $X(1)$ is easy to do in the vaccine group, but difficult in the placebo group as $X(1)$ is not observed. In this paper we derive bounds on the regression curve in the placebo group. For a continuous endpoint these bounds can be unhelpful, or can help modestly temper our enthusiasm for a role of $X(1)$ on the vaccine effect. For binary outcomes with 100% placebo infection the bound is very tight but unhelpful as 100% infection precludes identification of any covariate with a differential effect on placebo infection. We apply these methods to experiments of anthrax vaccine in rabbits with survival to challenge as the outcome and demonstrate how to extrapolate the model to humans.

KEYWORDS: causal, counterfactual, potential outcome, principal stratification

Author Notes: We thank the U.S. Army Research Institute of Infectious Disease for use of their data. We also thank Bob Kohberger, Freyja Lynn, and Ed Nuzum for helpful discussions.

1 Introduction

An area of intense interest in vaccine development is understanding the immune mechanisms that are responsible for protection (Qin et al 2007; Gilbert, Qin, Self 2008, Plotkin 2008). The temptingly simple approach of measuring vaccine induced immune responses in the vaccine group, say $X(1)$, and correlating $X(1)$ with protection or benefit can be misleading as $X(1)$ may be merely identifying those with a constitutional ability to remain uninfected and $X(1)$ may have no causal role. This seemed to be the likely conclusion from the VAX004 trial where an HIV vaccine with no overall benefit nonetheless showed a decreased infection rate as a function of $X(1)$ in the vaccine group (Gilbert et al 2005). Dissecting whether a vaccine induced immune response has a causative role or is merely identifying individuals who have good natural response is enormously important. Vaccine induced immune responses that show some glimmer of causation can, in principle, be improved by modifications that increase the causative immune responses. Vaccine induced immune responses that merely identify good responders may be interesting, but are not targets for vaccine development. Follmann (2006) introduced augmented vaccine designs that can address this concern but require either a baseline variable correlated with the immune response of interest, or the ability to vaccinate the placebo uninfecteds at the end of the study. At times, such maneuvers will not be possible and so other methods are needed to address this issue.

In this paper we develop bounds on the relationship between outcome and the potential outcome $X(1)$ in the placebo group. To aid understanding, we begin by developing this sensitivity model for a post-infection outcome such as viral load and then extend the model to a survival outcome. We conclude by analyzing data from experiments of an anthrax vaccine in rabbits, and use this data to impute what vaccine efficacy would have been in a trial of humans.

2 A post-infection outcome

Let $Z = 0, 1$ denote, respectively, the placebo and vaccine groups, Y be an indicator of protection (e.g. lack of infection or survival, depending on the case), and V be a post-infection outcome such as viral load. Denote by n_Z the number randomized and infected in group Z . Additionally, let $X(1)$ be the immune response that is measured shortly after vaccination in the vaccine group. In the placebo group, $X(1)$ represents the immune response they would have had, had they received the vaccine. Note that $X(1)$ is not observable in the placebo group and can thus be viewed as missing by design. However, we know that the distribution of $X(1)$ must

be the same in the placebo and vaccine groups by randomization. For completeness note that $X(0)$, the immune response a patient would have under placebo is always undetectable in our settings and not of interest in our development.

To proceed we will make several assumptions labeled C1, C2, ... for the continuous setting.

- **C1** The vaccine has no effect on acquisition of infection and everyone in both arms becomes infected. *This assumption is for simplicity and allows us to focus on key issues. Relaxation of C1 will be discussed later.*
- **C2** We assume that a linear regression holds between the potential viral load outcomes, the vaccine, and $X(1)$,

$$V(Z) = \beta_0 + \beta_1 Z + \beta_2 X(1) + \beta_3 ZX(1) + e(Z), \quad (1)$$

where $e(Z)$ is an error term with mean 0 and variance σ^2 . We assume that the two equations (for $Z = 0, 1$) apply to each person in the trial. The outcome $V(0)$ is the viral load a patient would have under placebo and $V(1)$ the viral load the same patient would have under vaccine.

Evaluating the parameters in (1) is instructive. Our primary interest is with the causal estimand

$$E\{V(1) - V(0)|X(1)\} = \beta_1 + \beta_3 X(1).$$

For example, if $\beta_3 \neq 0$ then the vaccine has an (vaccine - placebo) effect that varies with $X(1)$. For $X(1) = 0$, the effect is β_1 while for $X(1) = 3.1$ the effect is different, $\beta_1 + \beta_3 3.1$. However if $\beta_3 = 0$ and $\beta_1 \neq 0$ then the (vaccine - placebo) effect is constant at β_1 for all $X(1)$. If $\beta_3 = 0$ and $\beta_2 < 0$ then $X(1)$ is merely identifying individuals with a constitutional ability to lower viral loads and $X(1)$ is not predictive of the causal effect of the vaccine. Thus if $\beta_3 = 0$ there is no point in trying to increase $X(1)$ by vaccine modification. If $\beta_3 < 0$ then there is reason to hope that increasing $X(1)$ by vaccine modification might increase the causal effect of the vaccine. Unfortunately we can only estimate the vaccine slope $\beta_2 + \beta_3$ using standard methods.

If we saw $X(1)$ in everyone then (1) would be easy to estimate using standard regression. However, $X(1)$ is only observed in the vaccine group. A straightforward regression of $V(1)$ on $X(1)$ gives us a slope estimate of $\beta_2 + \beta_3$, but not the individual terms β_2, β_3 . So, for example, with a standard regression estimate of $\widehat{\beta_2 + \beta_3} = -3$ we cannot distinguish between e.g. $(\beta_2, \beta_3) = (-3, 0)$ or $(40, -43)$. Our goal is to develop methods that allow us to separately estimate β_2 and β_3 .

To proceed, we use the fact that with 100% infection in both groups, randomization implies that $f(X(1)|Z=0) = f(X(1)|Z=1) = f(X(1))$, the probability density function for $X(1)$. We next exploit a fact of least squares regression that any regression line of V on $X(1)$ must go through the sample averages $\overline{X(1)}, \overline{V}$. Let us define $\overline{X_Z(1)}, \overline{V_Z}$ as the sample averages of $X(1)$ and V in group $Z = 0, 1$. Note that $\overline{X_1(1)}, \overline{V_1}$ and $\overline{V_0}$ are directly observed. While $\overline{X_0(1)}$ is not seen in the placebo group, if n_0 and n_1 are large, it must be pretty close to the $\overline{X_1(1)}$ observed in the vaccine group by randomization. For clarity of exposition we ignore variability due to small n_Z . Consider Figure 1. In the vaccine group, we have an ordinary least squares regression. In the placebo group, we have different possibilities for the regression line though all must go through $(\overline{V_0}, \overline{X_0(1)})$. And $(\overline{V_0}, \overline{X_0(1)})$ must be pretty close to $(\overline{V_0}, \overline{X_1(1)})$. The family of possible placebo regression lines can thus be visualized by a spinner that is nailed down at the red square.

Since all placebo regression lines that go through $(\overline{V_0}, \overline{X_0(1)})$ are possible, we need to impose additional constraints to come up with useful bounds. The following two bounds should be defensible in a number of settings.

- **C3** $\beta_2 \leq 0$

If β_2 were positive then placebo volunteers who would have larger $X(1)$ s do have larger viral loads than those who would have smaller $X(1)$ s. This seems implausible.

- **C4** $\beta_3 \leq 0$

This ensures that the placebo line cannot be steeper than the vaccine line. Assumption C4 implies that the causal effect of the vaccine is nondecreasing in x .

The family of lines under assumptions C1-C4 lie between the two red dashed lines of Figure 1.

In some settings one might impose a bound consistent with the idea that the vaccine can never be harmful. Visually, this means that the dashed line never quite touches the solid line over the range of $X(1)$ observed in the vaccine group. Formally with the assumption of the linear model (1) we can specify this bound as

- **C5** $\beta_0 + \beta_1 + (\beta_2 + \beta_3)x \leq \beta_0 + \beta_2x$ for all x in the support of $X(1)$.

Figure 1: Regression relationships between viral load and immune response for some simulated data. Black vaccine data $(X_1(1), V_1)$ given by dots, red placebo data $(?, V_0)$ given by the dashes projected onto the Y axis. In the vaccine group, an ordinary least squares line is fit (solid line). In the placebo group, the regression line must nearly go through $(\bar{V}_0, \bar{X}_1(1))$, the red square. This restriction defines a family of lines. Two examples are given with $\beta_2 = 0$ and $\beta_3 = 0$.

Figure 2 provides a visual example of the bounds given by C3 and C5. If the two lines touch at the smallest $X(1)$, we can describe features of the vaccine effect over the range of $X(1)$. In this case, the estimated effect $\bar{V}_1 - \bar{V}_0 = \hat{\Delta}$ is also

the difference in predicted outcomes at $\overline{X(1)}$. For $X(1)$ symmetric about its mean, this bound provides an effect of $2\hat{\Delta}$ at the largest $X(1)$ and a bound of 0 at the smallest $X(1)$. The bound imposed by C5 depends critically on the assumption that $E(V(Z))$ is a linear function of $X(1)$ and thus behavior at the smallest $X(1)$ tells us something about what happens at the largest $X(1)$. In general, this kind of bound can be weakened. For example if we assume that $E(V(Z)|X(1))$ is a polynomial of $X(1)$ in both the placebo and vaccine group, we could have analogues to C3 (i.e. the placebo function $E(V(0)|X(1))$ is nonincreasing) and C5 (i.e. $E(V(1)|X(1) = x) \leq E(V(0)|X(1) = x)$ for all x in the support of $X(1)$). These bounds should be wider than that given by C3 and C5.

The bounds given by nonzero β_2 are most useful in deciding whether our enthusiasm for $X(1)$ should be tempered. For example, in Figure 1 there is about a 1.5 unit effect of the vaccine at the largest $X(1)$ compared to about a .25 unit effect at the smallest $X(1)$. If this is viewed as a modest effect, improvement of the vaccine may not be worth pursuing. Based on Figures 1 and 2 one can see the maximal effect will diminish as the mean viral load in the vaccine and placebo groups become more similar. Of course, one might have limited enthusiasm for a vaccine with a weak effect, period.

3 Infection endpoint

We can apply the same reasoning we used for $E(V(Z))$ to models for the probability of protection—a binary outcome. Here, however, there is no need to assume C1. As an analogue to (1) we assume the probability of protection follows

- **B1**

$$\begin{aligned} P(Y(Z) = 1|X(1), Z) &= \text{expit}(\beta_0 + \beta_1 Z + \beta_2 X(1) + \beta_3 X(1)Z) \\ &= \frac{1}{1 + \exp\{-\beta_0 - \beta_1 Z - \beta_2 X(1) - \beta_3 ZX(1)\}} \end{aligned} \quad (2)$$

In the vaccine world, vaccine effect is measured by the vaccine efficacy, VE, which is the proportion reduction in the infection rate in the vaccine group compared to the placebo group. At a specific $X(1)$ we have

$$VE\{X(1) = x\} = 1 - \frac{1 - P(Y = 1|X(1) = x, Z = 1)}{1 - P(Y = 1|X(1) = x, Z = 0)} \quad (3)$$

Figure 2: Regression relationships between viral load and immune response for some simulated data. Black vaccine data $(X_1(1), V_1)$ given by dots, red placebo data $(?, V_0)$ given by the dashes projected onto the Y axis. In the vaccine group, an ordinary least squares line is fit (solid line). In the placebo group, the regression line must nearly go through $(\bar{V}_0, \bar{X}_1(1))$, the red square. Two possibilities are given by the dashed lines; one where the effect of $X(1)$ is exactly the same in the vaccine and control groups ($\beta_3 = 0$), the other where we require that the vaccine is never harmful for any observed $X(1)$.

Again, $X(1)$ is missing by design in the placebo group but has the same distribution in both placebo and vaccine groups by randomization. Similar to the constraint of linear regression where the regression line must go through $\bar{V}_0(0), \bar{X}_0(0)$,

logistic regression maximum likelihood estimates (MLEs) are constrained so that the average estimated probability equals the overall success rate. Thus if we make the assumption that β_2 is fixed at A , then one can show that the MLE of β_0 solves the following equation

$$\bar{Y}_0 = \sum_{i=1}^{n_0} \frac{\text{expit}\{\beta_0 + AX_i(1)\}}{n_0}, \quad (4)$$

where $\bar{Y}_0$ is the mean protection rate in the placebo group. Of course the individual $X_i(1)$ are not observed in the placebo group so to solve (4) we need to come up with some specific values for the $X_i(1)$ s. We might take a set of $X_i(1)$ s from the vaccine group that are a representative or random sample.

Analogous to the linear regression setting, we can impose constraints on the slopes β_2 and β_3 which for the binary setting we denote as

- **B2** $\beta_2 \geq 0$
- **B3** $\beta_3 \geq 0$

The justification for these bounds parallels the arguments for the linear setting. The sign is different because larger values of the outcome now are better (i.e. survival) where as for the continuous case larger values were worse (i.e. more viremia). Interestingly, the infection rate in the placebo group actually has a profound effect on the utility of bounds for the placebo regression curve as we will see in the next section.

4 Anthrax vaccine experiments in rabbits

The US government has sponsored a series of experiments to evaluate the efficacy of vaccines to protect against anthrax. One vaccine under evaluation is known as AVA (Anthrax Vaccine Adsorbed), see Jollenbeck Zwanzinger Durch & Strom (2002). This vaccine is derived from culture filtrates from an avirulent anthrax strain. The filtrates contain antigens that are recognized by the immune system and provoke an adaptive immune response. AVA has a vaccination schedule of 5 shots at months 0, 1, 6, 12, and 18 and is licensed by the Food and Drug Administration. There is interest in reducing the number of shots required for this vaccine, and also to develop non-AVA vaccines that might require very few shots and have better immunogenicity and thus perhaps better protection. Towards this end, experiments have been performed in different animal species to see if an immune response can be identified that correlates with protection and hopefully has a causative role in protection.

If similar relationships between immune response and protection are seen in different animal species and if the VE varied with $X(1)$ one would be more comfortable using an animal model to infer the level of protection that could be achieved in man. In terms of the model (2), we are hoping to identify immune responses $X(1)$ with corresponding $\beta_3 > 0$.

As part of the extensive evaluation of this vaccine, experiments were conducted in 156 rabbits, with 18 rabbits receiving placebo and 138 receiving vaccine. Rabbits received injections at week 0 and week 4, immune responses were measured at weeks 6 and 10, and the animals were “challenged” at 10 weeks. Challenging rabbits is a euphemism for making rabbits inhale enough anthrax spores so that death is almost certain for unvaccinated rabbits.

The immune response of major interest is the maximum antibody response over weeks 6 and 10 as measured using a Toxin Neutralization Assay (TNA). This assay measures the ability of serum to prevent anthrax lethal toxin from killing cells *in vitro*. Briefly, fixed levels of lethal toxin and infectable cells are combined in wells along with serum in various dilutions. If there are many anthrax specific antibodies in the serum, a substantial dilution will still result in substantial protection of the susceptible cells. The dilution that achieves a 50% reduction in the cell killing (compared to no serum) or ED50 is estimated and we use the logarithm of this readout as our antibody measure.

To begin we fit an overall logistic regression model to the observed data Y_i, X_i , where Y_i is the survival indicator and X_i the measured immune response in both placebo and vaccine rabbits. In potential outcomes notation $X_i = Z_i X_i(1) + (1 - Z_i) X_i(0)$ and $Y_i = Z_i Y_i(1) + (1 - Z_i) Y_i(0)$. Figure 3 displays the results with both Y and X_i jittered to help visualization. Note that there is a striking relationship between measured immune response and survival. The slope and intercept are -12.26 and 2.12 with a Wald statistic for the slope of 5.85 . In the placebo group of 18 rabbits, all died and all had undetectable immune responses which are recorded as 3.26 .

In this paper we are interested in regression curves of survival on $X_i(1)$ separately in the placebo and vaccine groups. We directly fit a logistic regression model in the vaccine group:

$$P(Y(Z) = 1 | X(1), Z = 1) = \text{expit}\{(\beta_0 + \beta_1) + (\beta_2 + \beta_3)X(1)\}.$$

The results are virtually identical to the regression of Y on X_i which had 18 additional placebo rabbits who all died and had undetectable immune responses. In the placebo group $X(1)$ is not observed. However since all 18 placebo rabbits died it doesn't matter. More formally, for any postulated slope $\beta_2 = A$ and any set of $X_i(1)$ s

Figure 3: Estimated logistic regression curve for the probability of survival as a function of measured antibody response X . Black dots are vaccine data, red dots placebo data.

the solution to (4) is $\hat{\beta}_0 = -\infty$. Thus one simply estimates

$$\hat{P}(Y(Z) = 1|X(1), Z = 0) = 0$$

The fitted curves are displayed in Figure 4 along with the jittered data, with red being the placebo group and black the vaccine group. Visually one can see that this must be the only possible curve; if it is tilted at all, half the curve will be negative. Put another way, if some of the probabilities, $\text{expit}(\beta_0 + AX(1))$ of (4), are positive then others must be negative to satisfy $\bar{Y} = 0$. Since negative probabilities are impossible, the solution to (4) requires that all fitted probabilities must be zero.

Figure 4: Estimated logistic regression curve for the probability of survival as a function of the $X(1)$ antibody response which is measured in the vaccine group (black), but unobserved in the placebo group. Since all placebo rabbits die, the regression curve of survival on $X(1)$ in the placebo is 0 for all values of $X(1)$.

We estimate $VE\{X(1)\}$ as

$$VE\{X(1) = x\} = 1 - \frac{1 - \widehat{P}(Y = 1|X(1) = x, Z = 1)}{1 - \widehat{P}(Y = 1|X(1) = x, Z = 0)} \quad (5)$$

This actually simplifies for our data as the denominator is 1, but we want to develop the general case.

To assign uncertainty about the estimate of $VE\{X(1)\}$ we apply a parametric bootstrap to the numerator and denominator of the RHS of (5) separately. For the numerator we randomly generate Bernoulli random variables with probability $\widehat{P}(Y = 1|X_i(1), Z = 1)$ for $i = 1, \dots, 138$. We then obtain the estimates $\widehat{\beta}_0^* + \widehat{\beta}_1^*, \widehat{\beta}_2^* + \widehat{\beta}_3^*$ from the bootstrap sample and form

$$\widehat{P}^*(Y = 1|X(1) = x, Z = 1) = \frac{1}{1 + \exp(-(\widehat{\beta}_0^* + \widehat{\beta}_1^*) - (\widehat{\beta}_2^* + \widehat{\beta}_3^*)x)}.$$

For the denominator we randomly sample Bernoulli random variables with probability p_0 . For each placebo bootstrap sample we solve (4) for β_0 by fixing A at either 0 or $(\widehat{\beta}_2^* + \widehat{\beta}_3^*)$ —the estimated vaccine slope for the “matched” bootstrap sample from the vaccine group. For $X(1)$ s in (4) we randomly sample 18 $X(1)$ s from the vaccine group. For each bootstrap sample we then create two VEs corresponding to the two bounds given by B2 and B3.

$$VE^*(x)_{A=0} = 1 - \frac{\widehat{P}^*(Y = 1|X(1) = x, Z = 1)}{\widehat{P}_{A=0}^*(Y = 1|X(1) = x, Z = 0)},$$

and

$$VE^*(x)_{A=\widehat{\beta}_2 + \widehat{\beta}_3} = 1 - \frac{\widehat{P}^*(Y = 1|X(1) = x, Z = 1)}{\widehat{P}_{A=\widehat{\beta}_2 + \widehat{\beta}_3}^*(Y = 1|X(1) = x, Z = 0)}.$$

For our data set it is natural to use $p_0 = 0$ as all the placebo animals died. This makes the bootstrapping of the denominator pointless as we always estimate $\widehat{P}_A^*(Y = 1|X(1) = x, Z = 0) = 0$ for any A. This feels a little uncomfortable as the estimate of 0% placebo survival is only based on 18 animals, and we would get the same result for VE and its confidence interval if 0 of 18,000 animals survived. This is not so far off as other experiments in rabbits have also shown 0% placebo survival, so one might view it as based on a larger sample size. For illustration we will also repeat this exercise pretending that one of the placebo animals survived and use $p_0 = 1/18$ in the calculations of VE and the confidence intervals. We use the percentile method to form bootstrap confidence intervals. Table 1 presents the

Table 1

The estimated vaccine effect and 95% parametric bootstrap confidence intervals of the AVA vaccine in rabbits as a function of principle strata $X(1)$. We use both the actual data where 0/18 placebo animals survive and a perturbed dataset where 1/18 placebo animals survive.

Data	Principal Stratum Antibody Response $X(1)$	$\widehat{VE}\{X(d)\}$	95% Bootstrap percentile CI	
			$\beta_2 = 0$	$\beta_2 = \beta_2 + \beta_3$
Actual: No Placebo Survivors	6.0	.568	(.345,.708)	(.336,.719)
	7.5	.969	(.923,.995)	(.916,.995)
Made Up: One Placebo Survivor	6.0	.543	(.339, .719)	(.344, .729)
	7.5	.968	(.928,.996)	(.921, .995)

results at $X(1) = 6.0$ and 7.5 , values with moderate and high efficacy, respectively. We see very similar results for the two datasets with high efficacy for $X(1) = 7.5$.

Since we have a strong relationship between VE and $X(1)$, it seems that there should be more evidence that $X(1)$ has a causal effect on VE. But unfortunately, this is not really the case. With 100% lethality, it is not possible for the probability of survival to vary with $X(1)$ in the placebo group. Thus you will always conclude VE varies with $X(1)$ if the probability of survival varies with X in the vaccine group. Put another way, we have strong evidence that $\beta_2 + \beta_3 > 0$, but we cannot really tell if $\beta_3 = 0$ or not.

5 An estimate of the overall VE in a new setting

5.1 Extrapolating to an improved vaccine

In this section we apply the model developed in rabbits to humans in order to infer what VE might have been observed, had we conducted an unethical human challenge trial. Making a leap from a model in animals to a predicted VE in humans is substantial and tricky. To help elucidate key points for a more general setting where

β_2 is not equal to zero and $\beta_0, \beta_1, \beta_2, \beta_3$ are known, we will first discuss extrapolation from one vaccine to another for a continuous endpoint. For the continuous setting we assumed model (1) which implied a vaccine effect of

$$\beta_1 + \beta_3 X(1) \quad (6)$$

for those in the principal stratum $X(1)$. We will call this vaccine A. Now suppose that we have a reformulation, say vaccine B, which induces a better distribution of immune responses $X^B(1)$. We want to know what vaccine effect we can anticipate at a new principal stratum $X^B(1)$. For simplicity, assume that we have the deterministic relationship $X^B(1) = X(1) + \omega$. Thus the principal stratum $X(1) = X^A(1) = x$ is identical to the principal stratum $X^B(1) = x + \omega$. The mean response on placebo (within the experiment using either vaccine A or B) for individuals who have $X(1) = x, X^B(1) = x + \omega$ is

$$E(V^A(0)|X(1) = x) = E(V^B(0)|X^B(1) = x + \omega) = \beta_0 + \beta_2 x.$$

For the new vaccine we assume that

$$E(V^B(1)|X^B(1) = x + \omega) = \beta_0 + \beta_1 + \beta_2 x + \beta_3(x + \omega)$$

where $\beta_0, \beta_1, \beta_2, \beta_3$ apply to both vaccine A and B. Thus we anticipate the casual vaccine effect at $X^B(1) = x + \omega$ to be

$$E(V^B(1) - V^B(0)|X^B(1) = x + \omega) = \beta_1 + \beta_3(x + \omega).$$

We might have assumed this directly from (6), but this construction will be helpful as we now consider the more complicated setting of binary outcomes. Analogous to above, we assume that the probability of infection under vaccine B is given by

$$P(Y^B(0)|Z = 0, X^B(1) = x + \omega) = \text{expit}(\beta_0 + \beta_2 x),$$

and

$$P(Y^B(1)|Z = 1, X^B(1) = x + \omega) = \text{expit}(\beta_0 + \beta_1 + \beta_2 x + \beta_3(x + \omega)).$$

Again, where $\beta_0, \beta_1, \beta_2, \beta_3$ apply to both vaccine A and B. Thus for a hypothetical future trial of vaccine B with vaccine immune responses $X_1^B(1), \dots, X_N^B(1)$, we would predict the VE to be

$$\hat{V}E_{human} = 1 - \frac{1 - E\{1 + \exp(-\hat{\beta}_0 - \hat{\beta}_1 - \hat{\beta}_2 X + \hat{\beta}_3(X + \omega))\}^{-1}}{1 - E\{1 + \exp(-\hat{\beta}_0 - \hat{\beta}_2 X)\}^{-1}}.$$

where the expectation is with respect to the distribution of $X^A(1)$ (recall the strong assumption that $X^B(1) = X^A(1) + \omega$).

5.2 Extrapolating from rabbit to man

With this background, we can now consider extrapolating from rabbits to man. As part of the anthrax vaccine evaluation, humans have been immunized with AVA and their immune responses recorded, say $H_1(1), \dots, H_N(1)$. These are measured using the same method as the rabbits, but we use H rather than X to emphasize their humanity. We will estimate a VE (and its uncertainty) for a hypothetical vaccine trial in N humans who achieved immune responses $H_1(1), \dots, H_N(1)$. While we did observe a strong relationship with X and survival in the vaccine group, using our methods did not provide any additional evidence that this relationship was causal. Nonetheless, there is a substantial body of evidence that anthrax antibody responses are causal. In particular, passive transfer experiments have been done where anthrax specific antibodies have been cloned and injected into rabbits and other animals. Such animals show protection against lethal anthrax challenge (Migone et al 2009). We thus feel emboldened to apply the model estimated in rabbits to estimate an effect in immunized humans. Such extrapolation has been done before for anthrax.

We also assume that the model estimated in animals using $X(1)$ applies to humans with $H(1)$. Because of this assumption the subsequent development could be substantially simplified as we will get the same result no matter what β_2 is used and without loss could set it to zero. Thus we could say we have a bound at a single point of $\beta_2 = 0$. We keep in β_2 in our development to apply to settings where the model is based on say vaccine A and the extrapolation to vaccine B, as discussed earlier.

We assume that the probability of protection for a human given vaccine who achieves immune response $H(1) = h$ is given by

$$\hat{P}(Y_{human} = 1 | Z = 1, H(1) = h) = \frac{1}{1 + \exp\{-(\widehat{\beta_0 + \beta_1}) - (\widehat{\beta_2 + \beta_3})h\}}$$

where $(\widehat{\beta_0 + \beta_1}), (\widehat{\beta_2 + \beta_3})$ are estimated from vaccinated rabbits. To complete the calculation, we need to specify a model for the probability of protection for a human given placebo. If we directly use the model estimated from rabbits this is

$$\hat{P}(Y_{human} = 1 | Z = 0, H(1) = h) = 0. \quad (7)$$

However to be conservative, we can also fit models that assume the true placebo infection rate is 1/19 (recall all 18 placebo rabbits died) with either no relationship with $X(1)$ or where the placebo slope equals $(\widehat{\beta_2 + \beta_3})$. To be more formal, we postulate

$$\hat{P}(Y_{human} = 1 | Z = 0, H(1) = h) = \frac{1}{1 + \exp(-\beta_0 - \beta_2 h)}.$$

For the first setting we have $\beta_2 = 0$ and β_0 solves $1/\{1 + \exp(-\beta_0)\} = 1/19$. For the second setting $\beta_2 = (\beta_2 + \beta_3)$ and β_0 solves (4) with $\bar{Y}_P = 1/19$ with the $x_i^*(1)$ s of (4) corresponding to the observed human immune responses following vaccination, $H_1(1), \dots, H_N(1)$. The justification here is that we should have about a 1/19 death rate in a study of N new placebo rabbits=humans no matter what new vaccine the associated vaccinated rabbits=humans receive. And, of course, these N new placebo rabbits=humans would have potential immune responses similar to the $H_1(1), \dots, H_N(1)$ seen in the vaccinated humans=rabbits.

We thus estimate the VE in a human trial as

$$\hat{VE}_{human} = 1 - \frac{1 - \hat{P}_V}{1 - \hat{P}_P} \quad (8)$$

where

$$\hat{P}_V = \frac{1}{N} \sum_{i=1}^N \frac{1}{1 + \exp(-(\hat{\beta}_0 + \hat{\beta}_1) - (\hat{\beta}_2 + \hat{\beta}_3)H_i(1))} \quad (9)$$

and

$$\hat{P}_P = \frac{1}{N} \sum_{i=1}^N \frac{1}{1 + \exp(-\hat{\beta}_0 - \hat{\beta}_2 H_i(1))} \quad (10)$$

Note that $\hat{P}_P = 1/19$ for both scenarios: $\hat{\beta}_2 = 0$ and $\hat{\beta}_2 = (\hat{\beta}_2 + \hat{\beta}_3)$. In other words, unconditionally the overall placebo infection rate is 1/19 for either conditional model. This approach to estimating vaccine efficacy has been suggested by Williamson, Duchars & Kohberger (2010) among others.

To obtain an confidence interval for $\hat{VE}_{human}$ we apply the parametric bootstrap.

- Bootstrap the 138 vaccine rabbits: Draw a parametric bootstrap Y_i^* from Bernoulli $\hat{P}(Y = 1|Z = 1, X(1) = x_i)$ for $i = 1 \dots, 138$, where $\hat{P}$ is based on the original data.
- Bootstrap the 18 placebo rabbits: Draw Y_i^* from Bernoulli (1/19) for $i = 137, \dots, 156$.
- Estimate the parameters from the bootstrap sample. For simplicity we estimate $\beta_0, \beta_1, \beta_2, \beta_3$ subject to the constraint that $\beta_2 = 0$ (we obtain the same result if $\beta_2 = \beta_2 + \beta_3$). Denote these as $\hat{\beta}_0^*, \hat{\beta}_1^*, \hat{\beta}_2^*, \hat{\beta}_3^*$
- Apply the above estimated vaccine model to the human immunogenicity data $H_1(1), \dots, H_N(1)$ to estimate the VE in humans:

$$\widehat{VE}_{human}^* = 1 - \frac{1 - E\{1 + \exp(-(\hat{\beta}_0^* + \hat{\beta}_1^* - (\hat{\beta}_2^* + \hat{\beta}_3^*)H)\}^{-1}}{1 - E\{1 + \exp(-\hat{\beta}_0^* - \hat{\beta}_2^*H)\}^{-1}}$$

We repeat the above steps many times and form a percentile based bootstrap confidence interval for VE_{human} .

We apply this method to two AVA vaccine regimens; one the current 5 shot regimen at months 0, 1, 6, 12, and 18 and the other a 2 shot regimen at months 0 and 1. For both regimens, immune response was measured 4 weeks after the last two shots, and we take the maximum of these as our human immune response $H(1)$.

The 2 shot regimen was evaluated in 206 humans with a median $H(1)$ of 5.2. The overall VE is only .26 with a 95% confidence interval of (.10, .38). This analysis suggests that a 2 shot regimen is insufficient. If $X(1)$ has no effect on the placebo group (i.e., $\beta_2 = 0$) but some effect on the vaccine group (i.e., $\beta_3 > 0$), then modifying the vaccine regimen to increase the $X(1)$ and hence increase the VE could be possible. For example, giving 5 shots instead of two shots may increase the value of $X(1)$. In fact, when the 5 shot regimen was evaluated in 58 humans, it had a larger median $H(1)$ of 7.6. Under the assumption that $\beta_2 = 0$, then the overall VE for the 5 shot regimen is estimated as .92 with a 95% confidence interval of (.87, .96). Under this assumption, the vaccine administered to active US military personnel is estimated to have very high efficacy.

While the AVA vaccine has not been evaluated in a randomized placebo controlled trial, a similar formulation has been. Raw goat hair contains anthrax spores which can lead to infections both on the skin and in the lungs of humans who are exposed. Brachman et al (1962) review several studies in which textile mill workers who worked with raw goat hair were randomized to placebo or vaccine and then watched for instances of anthrax disease. The overall estimated vaccine efficacy was .925, very similar to the point estimate suggested by our analysis.

Another approach to estimate vaccine efficacy was suggested by Rubin (2004). Under this approach one generates potential outcomes $(Y(0)_i, Y(1)_i)$ $i = 1, \dots, N$ as Bernoulli random variables with probabilities given by (9) and (10). We do this many times, say M and then average the imputed VEs to get an estimate of the VE one would obtain in a trial of N humans. As $M \rightarrow \infty$, this average will converge to

$$1 - E\left(\frac{1 - \bar{Y}^*(1)}{1 - \bar{Y}^*(0)}\right) \quad (11)$$

where $\bar{Y}^*(1)$ is the mean of N Bernoullis with probabilities (9), $\bar{Y}^*(0)$ is the mean of N Bernoullis with probabilities (10), and E is with respect to the simulation distribution.

This estimate differs somewhat from the estimate (8) which can be written as

$$1 - \left(\frac{1 - E(\bar{Y}^*(1))}{1 - E(\bar{Y}^*(0))} \right) \quad (12)$$

As $N \rightarrow \infty$ (11) and (12) become identical.

6 Discussion

This paper has proposed a bounding approach to help estimate separate placebo and vaccine group regressions of outcome on an immune response $X(1)$ that is only observed in the vaccine group. This allows us to estimate the causal effect of vaccine within principal strata defined by $X(1)$. For continuous outcomes, these methods can help temper our enthusiasm for the role of an immune response for a vaccine with a modest effect. While we would probably have limited enthusiasm for any immune response induced by a vaccine with modest effect, our development allows for a formalism of this idea using the language of principal stratification. For a binary response the behavior of the bounds changes. In particular, with 100% placebo response, these methods will never lead to concluding that the vaccine effect is free of $X(1)$. The inherent problem is that with 100% placebo response, there is no possibility of identifying a covariate, counterfactual or otherwise, that has a relationship with response in the placebo group.

Another possible use of these bounds is in conjunction with other methods that help recover $X(1)$, such as the augmented designs of Follmann (2006). An obvious approach would be constrained maximum likelihood estimation of the placebo slope β_2 . As is generally true for constrained estimation, the resultant estimate should be biased away from the bound but would have smaller mean square error than an unconstrained estimate.

In our development of the continuous case we made the strong assumption that all subjects were infected. This assumption was made to help motivate our approach which was primarily focused on binary response in this paper. In practice, many subjects are often uninfected and the methods here would need to be modified. Under another strong assumption that the vaccine has no effect on acquisition, the bounds developed in this paper could be used directly. If the vaccine had an effect on acquisition one might develop new bounds perhaps using ideas in Gilbert

Bosch & Hudgens (2003) or Hudgens & Halloran (2006). Another tack would be to develop an overall likelihood that properly modeled the probability of infection as a function of $X(1)$ and other factors as mentioned in Follmann (2006).

REFERENCES

- Brachman PS, Gold H, Plotkin SA, Fekety R, Werrin M, Ingraham NR (1962) "Field evaluation of a human anthrax vaccine," *American Journal of Public Health and the Nation's Health* 52 632-646.
- Follmann DA (2006) "Augmented designs to assess immune response in vaccine trials" (2006) *Biometrics* 62 1161-1169.
- Gilbert PB, Peterson ML, Follmann D, Hudgens MG, Francis DP, Gurwith M, Heyward WL, Jobes DV, Popovic V, Self SG, Sinangil F, Burke D, Berman PW (2005) "Correlation between immunologic responses to a recombinant glycoprotein 120 vaccine and incidence of HIV-1 infection in a phase 3 HIV-1 preventive vaccine trial," *Journal of Infectious Diseases* 191 (5) 666-677.
- Gilbert PB, Bosch RJ, Hudgens MG (2003) "Sensitivity analysis for the assessment of causal vaccine effects on viral load in HIV vaccine trials," *Biometrics* 59 531-541.
- Gilbert PB, Qin L, Self SG (2008) "Evaluating a surrogate endpoint at three levels, with application to vaccine development," *Statistics in Medicine* 27, 4758-4778.
- Hudgens MG, Halloran ME (2006) "Causal vaccine effects on binary postinfection outcomes," *Journal of the American Statistical Association* 101 51-64.
- Jollenbeck LM, Zwanzinger LL, Durch JS, & Strom BL (2002) *The Anthrax Vaccine: Is it safe? Does it work?* The National Academies Press.
- Migone TS, Subramanian GM, Zhong J, Healey LM, Corey A, Devalaraja M, Lo L, Ullrich S, Zimmerman J, Chen A, Lewis M, Meister G, Gillum K, Sanford D, Mott J, Bolmer SD. (2009) "Raxibacumab for the treatment of inhalational anthrax" *New England Journal of Medicine* 361 135-44.
- Plotkin SA (2008) "Correlates of vaccine-induced immunity," *Clinical Infectious Diseases* 47, 401-409.
- Qin L, Gilbert PB, Corey L, McElrath, MJ, & Self SG (2007) "A framework for assessing immunological correlates of protection in vaccine trials," *Journal of Infectious Diseases* 196 1304-1312.

- Rubin DB (2004) ‘Direct and indirect causal effects via potential outcomes,’ *Scandinavian Journal of Statistics* 31 161-170.
- Williamson DE, Duchars MG, Kohberger R (2010) “Predictive models and correlates of protection for testing biodefence vaccine,” *Expert Review of Vaccines* 9 527-537.