Published in final edited form as:

Biometrics. 2018 March ; 74(1): 27–39. doi:10.1111/biom.12737.

Evaluating Principal Surrogate Markers in Vaccine Trials in the Presence of Multiphase Sampling

Ying Huang^{1,2,*}

¹Fred Hutchinson Cancer Research Center, Seattle, Washington, 98109, U.S.A

²Department of Biostatistics, University of Washington, Seattle, Washington, 98109, U.S.A

Summary

This paper focuses on the evaluation of vaccine-induced immune responses as principal surrogate markers for predicting a given vaccine's effect on the clinical endpoint of interest. To address the problem of missing potential outcomes under the principal surrogate framework, we can utilize baseline predictors of the immune biomarker(s) or vaccinate uninfected placebo recipients at the end of the trial and measure their immune biomarkers. Examples of good baseline predictors are baseline immune responses when subjects enrolled in the trial have been previously exposed to the same antigen, as in our motivating application of the Zostavax Efficacy and Safety Trial (ZEST). However, laboratory assays of these baseline predictors are expensive and therefore their subsampling among participants is commonly performed. In this paper we develop a methodology for estimating principal surrogate values in the presence of baseline predictor subsampling. Under a multiphase sampling framework, we propose a semiparametric pseudo-score estimator based on conditional likelihood and also develop several alternative semiparametric pseudo-score or estimated likelihood estimators. We derive corresponding asymptotic theories and analytic variance formulas for these estimators. Through extensive numeric studies, we demonstrate good finite sample performance of these estimators and the efficiency advantage of the proposed pseudo-score estimator in various sampling schemes. We illustrate the application of our proposed estimators using data from an immune biomarker study nested within the ZEST trial.

Keywords

Baseline Predictor; Closeout Placebo Vaccination; Estimated likelihood; Multiphase Sampling; Principal Surrogate; Pseudoscore

1. Introduction

The identification of immune surrogate markers plays an important role in vaccine research towards the prevention of infectious diseases (Plotkin, 2010). Since efficacy trials measuring clinical endpoints of interest are costly and operationally challenging to conduct, early vaccine development relies heavily on phase I–II immunogenicity studies in which candidate

* yhuang@fhcrc.org.

6. Supplementary Materials

Web Appendices referenced in Sections 2–5 are available with this paper at the *Biometrics* website on Wiley Online Library.

vaccines are screened and selected based on the magnitude, breadth, and quality of the immune responses they elicit. The identification of immune response markers (i.e. surrogate endpoints) that can reliably predict a vaccine's protective effect, is thus essential for reducing the cost and time of vaccine development.

This paper focuses on evaluating immune surrogates under the principal surrogate (PS) framework (Frangakis and Rubin, 2002), using data from a randomized trial. As discussed in Gilbert et al. (2015), this framework is particularly advantageous for surrogates evaluation in vaccine studies compared to alternative frameworks (Joffe and Greene, 2009) such as those based on assessing the Prentice criteria (Freedman et al., 1992; Lin et al., 1997; Burzykowski et al., 2006), direct and indirect effects of treatment (Robins and Greenland, 1992), and a meta-analytical approach that combines multiple trials (Daniels and Hughes, 1997; Buyse et al., 2000; Li et al., 2011). Under the PS framework, we can assess an immune response's surrogate value through the estimation of a vaccine efficacy (VE) curve, which contrasts clinical risks under each treatment assignment conditional on the vaccine-induced immune response (Gilbert and Hudgens, 2008). Examination of the VE curve helps inform the potential to improve a vaccine's efficacy by improving its immunogenicity.

The fundamental challenges in evaluating PS using VE curves, as a consequence of missing potential outcomes, are to ensure the validity of the estimator under non-identifiability issue and the precision of the estimator. To address missing potential outcomes, Follmann (2006) proposed the addition of two design components: the "baseline immunogenicity predictors" (BIP) component and a closeout placebo vaccination (CPV) component. The BIP component utilizes baseline predictors correlated with the unobserved immune biomarkers to predict the latter. It allows the identification of principal surrogate estimands under untestable assumptions on the risk model (Gilbert and Hudgens, 2008). The CPV component augments a standard trial design by vaccinating uninfected placebo recipients at the end of the trial and measuring their subsequent biomarker values. These values are then treated as if they had been recorded from subjects assigned to vaccine at the beginning of the trial, under additional assumptions that will be detailed below. The inclusion of the CPV component allows nonparametric estimation of disease risks under each treatment assignment, and consequently, the assessment of risk model assumptions (Huang and Gilbert, 2011).

When the BIP component is utilized, the precision of the principal surrogacy estimand depends critically on the strength of the correlation between the BIP and the potential biomarker (Follmann, 2006). Baseline covariates that are easy to acquire from participants (e.g., demographics) typically have low correlations with vaccine-induced immune responses and are therefore not useful as BIPs. However, biomarker measures related to the immune response of interest are potentially good baseline predictors. For example, in our motivating Zostavax Efficacy and Safety Trial (ZEST), baseline varicella zoster virus (VZV) specific immunity exhibited a good correlation with the candidate surrogate of interest, i.e. vaccine-induced change in VZV-specific immune response from baseline to week 6. In general, lab-assay-based BIPs are expensive to measure and thus are usually only acquired from a subset of the trial participants, as in the ZEST trial. This approach results naturally in a three-phase sampling design when the candidate surrogate (vaccine-induced immune response) is further measured among a subset of participants for whom the BIP is available.

While many statistical methods have been developed in recent years for evaluating PS under designs with BIP and/or CPV components, most of these methods focus on a two-phase sampling setting and assume that BIP is available from all trial participants whereas the potential surrogate is obtained from only a subsample of participants (Wolfson and Gilbert, 2010; Huang and Gilbert, 2011; Huang et al., 2013; Gabriel and Gilbert, 2014). Methods for dealing with BIP subsampling have not been rigorously studied or evaluated. Gilbert and Hudgens (2008) considered BIP subsampling and proposed fully parametric and nonparametric estimated likelihood estimators for design with the BIP component only. However, the asymptotic distributions of these estimators were not developed, the inference of which rely solely on resampling techniques. Moreover, the nonparametric estimated likelihood estimator developed in Gilbert and Hudgens (2008) requires the surrogate marker to be discrete. The research described here aims to fill this gap. In particular, we describe a methodology for evaluating continuous biomarkers as principal surrogates in the presence of BIP subsampling with or without a CPV component. Under a multiphase sampling framework, we propose a novel semiparametric pseudo-score estimator based on conditional likelihood and also investigate several alternative semiparametric estimators, including an estimated likelihood estimator based on conditional likelihood, and two other novel pseudo-score and estimated likelihood estimators based on weighted likelihood. For each estimator we develop algorithms based on weighted generalized linear model fitting that is easy to implement using standard statistical software. We derive analytical variance formulas for these estimators for making inference in practice and demonstrate through numerical studies the advantage of the proposed pseudo-score estimator over alternative estimators with respect to efficiency. Beyond vaccine trials, this research is anticipated to have broad application to general clinical trial surrogate endpoint evaluation and multiphase sampling problem.

In Section 2, we introduce the setting for evaluating principal surrogates, and describe the utility of the VE curve for quantifying surrogacy in vaccine trials. We then develop estimators for risk model parameters and for the VE curve. In Section 3, we evaluate and compare finite-sample performances of these estimators in various three-phase sampling designs. In Section 4 we illustrate the application of our proposed estimators using immunogenicity study data from the ZEST trial. We conclude with a discussion of the implications of our findings.

2. Method

We consider the setting of a two-arm randomized trial. Let Z be a binary indicator for treatment assignment, $Z = 0, 1$ for assignment to placebo and vaccine arm, respectively. Let Y be the binary disease outcome (infection status) observed at the end of the study, $Y = 0, 1$ for non-diseased and diseased, respectively. Let Q be baseline covariates used for modeling disease risk. We focus on discrete Q here, but note that the methods we describe can be generalized to accommodate continuous Q through the incorporation of nonparametric smoothing techniques. We assume Q can be partitioned into two components, X and W , with X available from all trial participants and W only available from a subset of the participants. For example, X might be a participant demographic such as age subgroup or gender; W might be a baseline laboratory measurement that is expensive to obtain. The special case of

$X = Q$ means that baseline covariates are measured from all trial participants, while $W = Q$ means that all baseline covariates used for risk modeling are subsampled. Let S be the candidate surrogate measured on the continuous scale at fixed time τ after randomization. We consider a univariate marker in this paper, although the estimation method we propose could be easily extended to allow for more than one marker. Let Y^τ be the indicator of whether disease develops before τ . S is only measurable if $Y^\tau = 0$ but undefined if $Y^\tau = 1$. We further incorporate the potential outcomes framework. Let $S(z)$, $Y^\tau(z)$, $Y(z)$ be the corresponding potential outcomes under treatment assignment z , for $z = 0, 1$.

In this paper we consider designs in which the set of subjects with S measured is among the set of subjects with W measured, as in our motivating ZEST example. In addition, we allow a possible CPV component in the study: at the end of the trial follow-up period, a fraction ($p_c \in [0, 1]$) of placebo recipients who are uninfected at study closeout and have W measured can be vaccinated with their immune biomarker S_c at time τ after vaccination measured. We call the design with no CPV (i.e. $p_c = 0$) the BIP-only design, and the design with non-zero CPV component (i.e. $p_c > 0$) the BIP + CPV design.

We characterize the problem under a three-phase sampling framework: 1) in the first phase, information of Y , Z , and X (if not empty) is available for all trial participants; 2) in the second phase, a subset of participants have W measured, with the sampling probability for W depending on Y , Z , and X ; 3) in the third phase, among those participants with W available, a subset have $S(1)$ or S_c measured, with the sampling probability for $S(1)$ and S_c depending on Y , Z , X , and W . Let δ_1 be a binary indicator for availability of W , and δ_2 be a binary indicator for availability of either $S(1)$ or S_c . For the settings we consider in this paper, $\delta_1 = 0$ implies $\delta_2 = 0$. Note that the special case of $\delta_1 = 1$ everywhere corresponds to the two-phase sampling problem where W is measured from all trial participants.

Our interest is to estimate the VE curve for the candidate surrogate, which is one type of the causal effect predictiveness surface (CEP) proposed by Gilbert and Hudgens (2008) for characterizing the principal surrogate effect. CEP contrasts the risk of $Y(1)$ and $Y(0)$ conditional on $S(1)$ and $S(0)$, i.e., $CEP\{S(1), S(0)\} = h[risk_{(1)}\{S(1), S(0)\}, risk_{(0)}\{S(1), S(0)\}]$ for a pre-specified contrast function h , with $risk_{(z)}\{S(1), S(0)\} = P\{Y(z) = 1 | S(1), S(0), Y^\tau(0) = Y^\tau(1) = 0\}$, $z = 0, 1$. A simplified version, the marginal causal effect predictiveness curve — $CEP\{S(1)\} = h[risk_{(1)}\{S(1)\}, risk_{(0)}\{S(1)\}]$ with $risk_{(z)}\{S(1)\} = P\{Y(z) = 1 | S(1), Y^\tau(0) = Y^\tau(1) = 0\}$ — is commonly used in vaccine research for the following reasons: In certain types of vaccine trials (e.g. HIV vaccine trials), only subjects without previous infection with the pathogen under study are enrolled such that $P\{S(0) = 0\} = 1$, and thus the task of estimating the CEP surface simplifies to the task of estimating the CEP curve. For other vaccine trials in which participants had previous exposure to the same pathogen, e.g. the ZEST trial, we can define the surrogate as the difference between post-vaccination immune response and baseline immune response, such that it is reasonable to consider $S(0)$ as scattered around zero with some noise due to measurement error, and the CEP curve is desirable (Gabriel and Gilbert, 2014). In vaccine research, a CEP curve of common interest is the VE curve. Let

$$VE\{S(1)\} = 1 - risk_{(1)}\{S(1)\} / risk_{(0)}\{S(1)\}, \quad (1)$$

i.e. the percent reduction in infection rate for the subgroup of vaccine recipients with immune response $S(1)$ compared to if they had not been vaccinated. The VE curve (curve of $VE(s)$ versus s) tells the range of VE potentially achievable for various levels of $S(1)$. A surrogate with a large variability in $VE\{S(1)\}$ is desirable due to a high potential to select efficacious vaccine regimens based on their immunogenicity in phase I/II trials. We estimate the VE curve by first estimating the disease risk conditional on $S(1)$ and Z , as detailed below.

2.1 Identifiability Assumptions

We next lay out a list of assumptions needed that have been commonly made in the literature (Gilbert and Hudgens, 2008; Huang and Gilbert, 2011; Huang et al., 2013).

- (A1) SUTVA and Consistency: $\{S(1), S(0), Y^\tau(1), Y^\tau(0), Y(1), Y(0)\}$ of one subject is independent of the treatment assignments of other subjects. Moreover, given the treatment a subject actually received, a subject's potential outcomes equal the observed outcomes.
- (A2) Ignorable Treatment Assignments: $Z \perp W, X, S(1), S(0), Y^\tau(1), Y^\tau(0), Y(1), Y(0)$.
- (A3) Equal Early Clinical Risk: $P\{Y^\tau(1) = Y^\tau(0) = 1\}$ for all subjects, which is reasonable if the biomarkers are measured shortly after randomization. Under (A3), the risk of Y conditional on $Z, W, X, S(1)$, and $S(0)$, i.e. $P\{Y(z) = 1 | S(1), S(0), W, X, Y^\tau(0) = Y^\tau(1) = 0\}$ for $z = 0$ or 1 , equals $P\{Y(z) = 1 | S(1), S(0), W, X, Y^\tau = 0\}$. Henceforth we simplify the notation and drop the conditioning of all probabilities on $Y^\tau(1) = Y^\tau(0) = Y^\tau = 0$.
- (A4) The risk of Y conditional on $Z, S(1), W$, and X follows a generalized linear model (GLM): $risk_{(Z)}\{S(1), W, X\} \equiv P\{Y(z) = 1 | S(1), W, X\} = g\{\beta; S(1), Z, W, X\}$, with g a pre-specified link function and β a finite-dimensional parameter.
- (A5) Time-constancy of immune response: for uninfected placebo recipients, $S(1) = S^{true} + U_1$, and $S_c = S^{true} + U_2$, for some underlying S^{true} and i.i.d. measurement error U_1, U_2 .
- (A6) No placebo subjects uninfected at closeout have an infection over the next τ time-units.

Assumptions (A5) and (A6) allow the use of S_c to substitute $S(1)$ for participants sampled in CPV. For the rest of the paper, we simplify the notation and use S to indicate a measurement of a vaccine-induced immune response that can be obtained either during the standard trial period or during CPV. Let N be the number of trial participants. The observed data are N iid copies $O_i = (Z_i, X_i, \delta_{1i}, \delta_{1i}W_i, \delta_{2i}, \delta_{2i}S_i, Y_i)'$, $i = 1, \dots, N$.

Next, we will introduce several semiparametric estimators of risk model and the VE curve under the three-phase sampling design. These estimators are semiparametric in the sense that they are based on a parametric functional form of the risk model (A4) for ease of interpretation, but rely on empirically estimated distributions of S and W for flexibility.

2.2 Estimators Based on Conditional Likelihood

The estimators we will describe in this section are based on the likelihood of infection outcome Y conditional on observed predictors of Z , X , W , and S , i.e.,

$$L(\beta) = \prod_{\delta_{1i}=\delta_{2i}=1} P(Y_i|Z_i, S_i, W_i, X_i) \prod_{\delta_{1i}=1, \delta_{2i}=0} P(Y_i|Z_i, W_i, X_i) \prod_{\delta_{1i}=\delta_{2i}=0} P(Y_i|X_i, Z_i) \\ = \prod_{\delta_{2i}=1} P(Y_i|Z_i, S_i, W_i, X_i) \prod_{\delta_{1i}=1, \delta_{2i}=0} \int P(Y_i|s, Z_i, W_i, X_i) dF(s|Y_i, W_i, X_i) \times \prod_{\delta_{1i}=0} \int P(Y_i|s, w, X_i) dF(s, w|X_i). \quad (2)$$

The corresponding score equation is

$$U = \sum_{\delta_{2i}=1} U_{\beta}(Y_i|Z_i, S_i, W_i, X_i) + \sum_{\delta_{1i}=1, \delta_{2i}=0} \frac{\int U_{\beta}(Y_i|s, Z_i, W_i, X_i) P(Y_i|s, Z_i, W_i) dF(s|W_i, X_i)}{\int P(Y_i|s, Z_i, W_i) dF(s|W_i, X_i)} \\ + \sum_{\delta_{1i}=0} \frac{\int U_{\beta}(Y_i|Z_i, s, w) P(Y_i|s, w, Z_i, X_i) dF(s, w|X_i)}{\int P(Y_i|Z_i, X_i, s, w) dF(s, w|X_i)}, \quad (3)$$

where $U_{\beta}(Y_i|Z_i, S_i, W_i, X_i) = \log P(Y_i|Z_i, S_i, W_i, X_i) / \beta$. We next present two estimators for β as solutions to estimates of (3): our proposed pseudo-score estimator and an alternative estimated likelihood estimator.

2.2.1 Three-Phase Pseudo-score estimator (proposed)—Here we propose an estimator based on solving a pseudo-score derived from (3), which we name the three-phase pseudo-score estimator (TPS). Pseudo-score type estimators for two-phase sampling were first introduced in Chatterjee et al. (2003) and later refined for improved efficiency in randomized trials (Huang et al., 2013). Here we develop the TPS estimator for handling subsampling of W .

First, by Bayes' theorem, we have

$$dF(s|w, x) = \frac{dF(s|w, x, \delta_2=1)P(\delta_2=1|w, x)}{P(\delta_2=1|s, w, x)}, \quad (4)$$

$$dF(s, w|x) = \frac{dF(s, w|x, \delta_2=1)P(\delta_2=1|x)}{P(\delta_2=1|s, w, x)}. \quad (5)$$

Replacing $dF(s|w, x)$ in (3) with the right hand side (RHS) of (4) and replacing $dF(s, w|x)$ in (3) with the RHS of (5) yields the following equation:

$$U = \sum_{\delta_{2i}=1} U_{\beta}(Y_i|Z_i, S_i, W_i, X_i) + \sum_{\delta_{1i}=1, \delta_{2i}=0} \frac{\int U_{\beta}(Y_i|s, Z_i, W_i, X_i) \frac{P(Y_i|Z_i, s, W_i, X_i)}{P(\delta_2=1|s, W_i, X_i)} dF(s|W_i, X_i, \delta_2=1)}{\int \frac{P(Y_i|Z_i, s, W_i, X_i)}{P(\delta_2=1|s, W_i, X_i)} dF(s|W_i, X_i, \delta_2=1)} \\ + \sum_{\delta_{1i}=0} \frac{\int U_{\beta}(Y_i|Z_i, s, w, X_i) \frac{P(Y_i|Z_i, s, w, X_i)}{P(\delta_2=1|s, w, X_i)} dF(s, w|X_i, \delta_2=1)}{\int \frac{P(Y_i|Z_i, s, w, X_i)}{P(\delta_2=1|s, w, X_i)} dF(s, w|X_i, \delta_2=1)}. \quad (6)$$

Also we note that

$$P(\delta_2=1|s, W_i, X_i) = P(\delta_1=\delta_2=1|s, W_i, X_i) \\ = \sum_{z=0}^1 \sum_{y=0}^1 P(\delta_2=1|\delta_1=1, y, z, W_i, X_i) P(\delta_1=1|y, z, X_i) P(y|s, z, W_i, X_i) P(z). \quad (7)$$

To compute the TPS estimator, we estimate $R(w|x, \delta_2=1)$ empirically with $\Sigma_{\delta_{2i}=1} I(W_i = w, X_i = x) / \Sigma_{\delta_{2i}=1} I(X_i = x)$ for each unique value x of X , estimate $R(s|w, x, \delta_2=1)$ empirically with $\Sigma_{\delta_{2i}=1} I(S_i = s, W_i = w, X_i = x) / \Sigma_{\delta_{2i}=1} I(W_i = w, X_i = x)$ for each unique combination of w and x , and estimate $P(\delta_1=1|y, z, x)$ and $P(\delta_2=1|\delta_1=1, y, z, w, x)$ in (7) nonparametrically. Entering these into (6) leads to a “pseudo-score” equation. The TPS estimator of β is then obtained as the solution to this equation. We solve for β using an algorithm that repetitively fits a weighted GLM as below:

- I. Start with an initial value of β ;
- II. (a) For a subject i with $\delta_{1i} = \delta_{2i} = 1$, use its observed data. (b) For a subject with $\delta_{1i} = 1$ and $\delta_{2i} = 0$, construct a set of filled-in data with length equal to the number of observations in $\mathcal{V}_{W_i, X_i}$ where $\mathcal{V}_{W_i, X_i}$ is the set of validating observations with $\delta_{2i} = 1$, $W = W_i$ and $X = X_i$. Specifically, for each $j \in \mathcal{V}_{W_i, X_i}$ we construct a new observation $\{Y_j, S_j, Z_j, W_j, X_j\}$. (c) For a subject with $\delta_{1i} = 0$ and $X = X_i$, construct a set of filled-in data with length m equal to the total number of observations in $\mathcal{V}_{W_k, X_i}$ across $k = 1, \dots, K$, i.e. $m = \sum_{k=1}^K \sum_j I(\delta_{2j}=1, W_j=k, X_j=X_i)$. Specifically, for each $j \in \mathcal{V}_{W_k, X_i}$ we construct a new observation $\{Y_j, S_j, Z_j, W_k, X_j\}$, and we repeat this for each $k = 1, \dots, K$.

- III. For a subject i with $\delta_{1i} = \delta_{2i} = 1$, give it weight 1. For each filled-in observation in II(b), $\{Y_i, S_j, Z_i, W_i, X_i\}, j \in \mathcal{V}_{W_i, X_i}$ calculate an associated weight,

$$\frac{\hat{P}(Y_i|S_j, Z_i, W_i, X_i)/\hat{P}(\delta_1=\delta_2=1|S_j, W_i, X_i)}{\sum_{j \in \mathcal{V}_{W_i, X_i}} \hat{P}(Y_i|S_j, Z_i, W_i, X_i)/\hat{P}(\delta_1=\delta_2=1|S_j, W_i, X_i)},$$

where $\hat{P}(\delta_1 = \delta_2 = 1|S_j, W_i, X_i)$ is estimated using (7) with $\hat{P}(\delta_2 = 1|\delta_1 = 1, y, z, W_i, X_i)$, $\hat{P}(\delta_1 = 1|y, z, W_i, X_i)$, and $\hat{P}(y|S_j, z, W_i, X_i)$ substituted for their true values, with $\hat{P}(y|S_j, z, W_i, X_i)$ obtained based on the current β estimate. For each filled-in observation in II(c), $\{Y_i, S_j, Z_i, w_k, X_i\}, j \in \mathcal{V}_{w_k, X_i}, k = 1, \dots, K$, calculate an associated weight,

$$\frac{\hat{P}(Y_i|S_j, Z_i, w_k, X_i)/\hat{P}(\delta_1=\delta_2=1|S_j, w_k, X_i) \times d\hat{F}(w_k|X_i, \delta_2=1)}{\sum_{k=1}^K \sum_{j \in \mathcal{V}_{w_k, X_i}} \hat{P}(Y_i|S_j, Z_i, w_k, X_i)/\hat{P}(\delta_1=\delta_2=1|S_j, w_k, X_i) \times d\hat{F}(w_k|X_i, \delta_2=1)}.$$

- IV. Fit a weighted GLM to the augmented dataset and obtain a new estimate of β .
- V. Repeat steps (I) to (IV) until convergence.

Under the regularity conditions specified in Web Supplementary Appendix A1, the TPS estimator can be shown to be consistent and asymptotically normally distributed. Theorem 1 presented in Supplementary Appendix A2 describes the asymptotic distribution of $\hat{\beta}$, with a proof outlined in Supplementary Appendix A3.

Later in our simulation studies and data example, we consider a risk model $P\{Y=1|S(1), Z, W\} = \Phi\{\beta_0 + \beta_1 Z + \beta_2 S(1) + \beta_3 S(1)Z\}$. Based on TPS estimators $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$, we estimate $VE(s)$ with $\widehat{VE}(s) = 1 - \Phi\{\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 s + \hat{\beta}_3 s\} / \Phi\{\hat{\beta}_0 + \hat{\beta}_2 s\}$. Asymptotic normality of $\widehat{VE}(s)$ follows from Theorem 1 based on the Delta method. Its asymptotic variances can be derived as: $\text{var}[\widehat{VE}(s)] = \left[\partial VE(s) / \partial \beta|_{\hat{\beta}} \right]^T \text{var}(\hat{\beta}) \left[\partial VE(s) / \partial \beta|_{\hat{\beta}} \right]$.

In the sampling design we consider here, the true sampling probabilities for W and S , i.e. $P(\delta_1 = 1|y, z, x)$ and $P(\delta_2 = 1|\delta_1, y, z, x, w)$, are known and could have been used in deriving the TPS estimator. In many two-phase sampling problems, however, using the “estimated” sampling weight has been recommended even when the true sampling weight is known because of the efficiency gain (Robins et al., 1994). Through extensive numerical studies in our multiphase sampling setting, estimating the sampling weight in general has comparable efficiency relative to using the true sampling weight and in some scenarios improves efficiency. We thus adopt the estimated sampling probability in our TPS estimator in this paper, and also for the several alternative estimators we will introduce next.

2.2.2 Three-phase Estimated Likelihood Estimator—In this section, we consider an alternative estimated-likelihood type (Pepe and Fleming, 1991) estimator based on maximizing an estimated conditional likelihood (2) by solving the corresponding estimated

score equation (3). We name this estimator the three-phase estimated likelihood estimator (TEL).

Unlike in TPS where we estimate distribution of S and W conditional on X among those with both S and W sampled ($\delta_2 = 1$), here we directly estimate the conditional distribution of S and W in the population and account for biased sampling of these variables using inverse probability weighting (IPW) (Horvitz and Thompson, 1952). In particular, we estimate $R(w|x)$ nonparametrically using data from all individuals with W measured, with the contribution of each individual weighted by the inverse of his/her probability of sampling W , i.e.

$$\hat{F}(w|x) = \frac{\sum_i I(W_i \leq w, X_i = x) / \hat{P}(\delta_1 = 1 | Y_i, X_i, Z_i)}{\sum_i I(X_i = x) / \hat{P}(\delta_1 = 1 | Y_i, X_i, Z_i)}$$
. For estimation of $R(s|w, x) = R(s|Z=1, w, x)$, we use data from vaccine recipients who have both S and W measured, with inverse probability weighting used to account for biased sampling of S , i.e.,

$$\hat{F}(s|w, x) = \frac{\sum_i I(S_i \leq s, W_i = w, X_i = x, Z_i = 1) / \hat{P}(\delta_2 = 1 | Y_i, X_i, W_i, Z_i)}{\sum_i I(W_i = w, X_i = x, Z_i = 1) / \hat{P}(\delta_2 = 1 | Y_i, X_i, W_i, Z_i)}$$
. Note that placebo recipients who have $\delta_2 = 1$ can not contribute to estimation of $R(s|w, x)$ since probability of obtaining S for placebo recipients who are infected at the end of the standard trial followup equals zero and thus IPW is not applicable. Substituting the above estimates of $R(s|w, x)$ and $R(w|x)$ into (2) and (3), we obtain the estimated conditional likelihood and corresponding estimated score equation. The TEL estimator of β is then derived as the maximizer of the estimated likelihood or the solution of the estimated score. Note that estimated likelihood estimators of a similar flavor have been previously proposed assuming a parametric distribution of S conditional on W or for discrete S (Gilbert and Hudgens, 2008). In Supplementary Appendix B, we propose an algorithm to solve for β by repetitively fitting a weighted GLM; we show that the TEL estimator of β is asymptotically normally distributed and provide the analytical formula for its variance. The asymptotic normality for TEL estimator of $VE(s)$ follows.

While TPS and TEL both solve for an estimated version of the score equation (3), TPS transforms the task of estimating the distribution of S , W conditional on X in the population as in TEL to the estimation of their distributions among samples with S and W measured ($\delta_2 = 1$). As mentioned earlier, when a CPV component is present, placebo recipients in CPV will contribute to the estimation of the conditional distribution of S in TPS but not in TEL. In earlier research under a two-phase sampling setting, a pseudo-score estimator has been shown to be more efficient compared to an estimated likelihood estimator in a BIP+CPV design (Huang et al., 2013). The same observation regarding the comparison between TPS and TEL is made later in our simulation studies. Another difference between TPS and TEL that is unique to the three-phase sampling problem regards the use of samples for estimating the conditional distribution of W : in TPS, only samples with $\delta_2 = 1$ contribute to the estimation of $R(W|X, \delta_2 = 1)$, whereas in TEL, all samples with $\delta_1 = 1$ contribute to the estimation of $R(W|X)$. While we might speculate that this will lead to efficiency loss of TPS relative to TEL in certain BIP-only designs in which a large proportion of samples have W measured but not S , the impact on estimation of the VE curve appears to be minimal as demonstrated by various simulation studies we conduct later.

2.3 Estimators based on weighted likelihood

In this section, we briefly describe two other new estimators derived from weighted likelihood that use inverse probability weighting to account for subsampling of W .

Consider a hypothetical scenario with W available from every trial participant. The likelihood of Y conditional on the observed predictors is $\prod_{\delta_2=1} P(Y_i|Z_i, W_i, X_i, S_i) \prod_{\delta_2=0} P(Y_i|Z_i, W_i, X_i)$. When W is subsampled, a consistent estimate of this likelihood can be constructed as a weighted likelihood using samples with W available: $L(\beta) = \prod_{\delta_1 \neq \delta_2=1} P(Y_i|Z_i, S_i, W_i, X_i)^{v_{1i}} \prod_{\delta_1=1, \delta_2 \neq 0} P(Y_i|Z_i, W_i, X_i)^{v_{1i}}$, where the contribution of each observation is weighted by the inverse of its sampling probability of W , i.e., $v_{1i} = 1/P(\delta_1 = 1|Y_i, Z_i, X_i, S_i) = 1/P(\delta_1 = 1|Y_i, Z_i, X_i, W_i, S_i)$. The corresponding score equation is

$$U = \sum_{\delta_{2i}=1} v_{1i} U_{\beta}(Y_i|Z_i, S_i, W_i, X_i) + \sum_{\delta_{1i}=1, \delta_{2i}=0} v_{1i} \frac{\int U_{\beta}(Y_i|s, Z_i, W_i, X_i) P(Y_i|s, Z_i, W_i, X_i) dF(s|W_i, X_i)}{\int P(Y_i|s, Z_i, W_i, X_i) dF(s|W_i, X_i)}. \quad (8)$$

As for TPS and TEL, we can construct a weighted pseudo-score and a weighted estimated likelihood estimators by solving an estimated version of the weighted score equation (8).

First, replacing $dF(s|w, x)$ with the RHS of (4), we can re-write (8) as

$$U = \sum_{\delta_{1i}=\delta_{2i}=1} v_{1i} U_{\beta}(Y_i|Z_i, S_i, W_i, X_i) + \sum_{\delta_{1i}=1, \delta_{2i}=0} v_{1i} \frac{\int U_{\beta}(Y_i|s, Z_i, W_i, X_i) \frac{P(Y_i|s, Z_i, W_i, X_i)}{P(\delta_1=\delta_2=1|s, W_i, X_i)} dF(s|W_i, X_i, \delta_1=\delta_2=1)}{\int \frac{P(Y_i|s, Z_i, W_i, X_i)}{P(\delta_1=\delta_2=1|s, W_i, X_i)} dF(s|W_i, X_i, \delta_1=\delta_2=1)} \quad (9)$$

We develop a weighted pseudo-score estimator (WPS) based on (9). In particular, we estimate $P(\delta_1 = 1|Y, Z, X)$, $P(\delta_2 = 1|\delta_1 = 1, Y, Z, W, X)$, and $P(s|W, Z, \delta_2 = 1)$ nonparametrically as in TPS. Substituting these into (9) and utilizing the relationship (7) results in a weighted pseudo-score. We propose to estimate β as the solution to the estimated pseudo-score by repetitively fitting a weighted GLM (Supplementary Appendix C1).

Alternatively, we can construct a weighted estimated likelihood estimator (WEL). In particular, we estimate $P(\delta_1 = 1|y, z, x)$ and $P(\delta_2 = 1|\delta_1 = 1, y, z, w, x)$ nonparametrically as before, and estimate $P(s|w, x) = P(s|Z = 1, w, x)$ with the IPW estimator as in TEL using vaccine recipients for whom both S and W have been measured. Substituting these into (8), we propose an algorithm that solves for β as the solution of the equation by repetitively fitting a weighted GLM (Supplementary Appendix D1). WPS and WEL are each asymptotically normally distributed as detailed in Web Supplementary Appendix C2 & D2, which show the derived analytical formulas for their variance.

3. Simulation Study

In this section, we evaluate the finite-sample performance of the proposed TPS estimator and of various comparative estimators including TEL, WPS, and WEL. We simulate S from a normal distribution with mean 3 and variance 1, and simulate a categorical W with four levels derived from discretizing an underlying normal random variable W^* correlated with S by quartiles. We consider settings with correlation between W^* and S being $\rho_{SW^*} = \{0.3, 0.5, 0.7\}$, which correspond to correlation $\{0.277, 0.462, 0.648\}$ between W and S . We assume a probit model for the risk of the binary outcome Y conditional on S , Z and W : $P(Y=1|S, Z, W) = \Phi(\beta_0 + \beta_1 Z + \beta_2 S + \beta_3 SZ)$. The risk model parameters are chosen such that the probability of infection is 0.12 and 0.06 in the $Z=0$ and $Z=1$ arms, respectively.

We consider a three-phase sampling design. In phase 1, $N=4,000$ subjects are randomized in 1:1 ratio of vaccine ($Z=1$) to placebo ($Z=0$). In this phase, information about Z and Y is available from every trial participant. In phase 2, stratified Bernoulli sampling of W is conducted: W is measured from all cases (infected participants) and a portion of controls (uninfected participants) sampled from the vaccine and placebo arms, respectively. In phase 3, stratified Bernoulli sampling of S is conducted among the participants for whom W is available: S is measured for all cases in the vaccine arm, and for a portion of controls in the vaccine arm or placebo arm (the CPV component). The performances of the different estimators are then compared under various sampling schemes. Each sampling scheme is characterized by four design parameters: i) γ_{w_v} , the ratio of the average number of vaccinated controls for whom W has been measured to the average number of vaccinated cases, ii) γ_{s_v} , the ratio of the average number of vaccinated controls for whom S has been measured to the average number of vaccinated cases, iii) γ_{w_p} , the ratio of the average number of placebo controls for whom W has been measured to the average number of vaccinated cases, and iv) γ_{s_p} , the ratio of the average number of placebo controls for whom S has been measured to the average number of vaccinated cases. For each scenario, results are presented based on 5,000 Monte-Carlo simulations. We evaluate performance for estimating $\beta_0, \beta_1, \beta_2, \beta_3$ and for estimating $VE(s)$ using equation (1) for s corresponding to various percentiles of $S(1)$.

First, we present finite sample performance for the four estimators of β and $VE(s)$ for various combinations of $\gamma_{w_v}, \gamma_{s_v}, \gamma_{w_p}$ and γ_{s_p} and for $\rho_{SW^*} = 0.5$. For BIP-only designs, Table 1 provides bias, variance, and coverage of 95% Wald confidence intervals (CI)

assuming approximate normality of $\hat{\beta}$ and $\log\{1 - \widehat{VE}(s)\}$ based on analytical variance estimates. All four estimators of β and $VE(s)$ have minimal biases in all settings, and their 95% Wald CIs have accurate coverage in general. For BIP+CPV designs, we observe the same good finite-sample performance for estimators based on conditional likelihood (TPS and TEL), in terms of small bias and good coverage probability (Table 2). Compared to TPS and TEL, the weighted likelihood based estimators (WPS and WEL) tend to be more variable and have larger bias for estimating $VE(s)$ when sample size is small (e.g. $\gamma_{w_v} = \gamma_{s_v} = \gamma_{w_p} = \gamma_{s_p} = 1$). However, they have satisfactory performance when sample size is larger.

We further compared the efficiency in estimating β and $VE(s)$ for various sampling schemes using either the BIP-only or the BIP+CPV design. Table 3 presents the efficiency of TPS

relative to each alternative estimator for estimating β . For BIP-only designs, the performances of WPS and WEL are similar; the estimator with the best efficiency varies with the coefficient of interest and the sampling scheme. For example, when estimating the intercept β_0 and main marker effect β_2 , TEL exhibits the best efficiency in some settings, whereas WPS and WEL perform best in others. However, TPS and TEL generally perform the best when estimating the main treatment effect β_1 or the interaction term β_3 . Note that even in settings when TPS and TEL are not the best performing methods, their efficiencies are typically fairly close to those of the best performing ones; in contrast, WPS and WEL exhibit appreciable efficiency loss compared to the best performing estimator in some settings. Also presented in Table 3 are efficiency comparisons for estimating β in BIP+CPV designs. In this design the TPS estimator performs best for estimating all β 's across all sampling schemes studied. Table 4 presents the efficiency of TPS relative to alternative estimators for estimating $VE(s)$ for s values corresponding to 10%, 30%, 50%, 70%, 90% percentiles of $S(1)$. For BIP-only designs, TPS and TEL in general perform better or similarly to WPS and WEL; for BIP+CPV designs, TPS always performs best and demonstrates a quite appreciable efficiency gain relative to other estimators in some settings.

Similar patterns for comparison between estimators are observed when we vary correlation between S and W^* to be 0.3 and 0.7 (Web Supplementary Tables 2–9).

4. Data Example

In this section, we apply our developed estimators for surrogacy to data from an immune response study nested within the Zostavax Efficacy and Safety Trial (ZEST), a phase III vaccine trial for studying the effect of the Zostavax vaccine against varicella zoster virus (VZV). In ZEST, 22439 North American and European subjects aged 50–59 with a history of varicella or at least 30 years residence in a VZV-endemic area were 1:1 randomly assigned to receive a single dose of zoster vaccine or placebo. Subjects were followed for an average of 1.3 years post-vaccination for the development of herpes zoster (HZ). Vaccine efficacy was estimated to be 69.8% (95% CI, 54%–81%) in ZEST (Schmader et al., 2012).

An immune correlates study was subsequently conducted among participants from the ZEST trial. In 10% of randomly selected study participants and in all cases who experienced HZ after the week 6 visit (25 cases in the vaccine arm and 91 cases in the placebo arm), VZV-specific antibody responses were measured using gpELISA on day 1 (before immunization, baseline) and week 6 (after immunization). A correlation of -0.57 was found between log-transformed baseline gpELISA and its fold rise at week 6 relative to baseline among vaccine recipients. Fold rise in gpELISA at week 6 relative to baseline was found to be an good correlate for predicting efficacy of herpes zoster vaccine (Gilbert et al., 2014).

In the following, we apply methods developed in this paper to these immune correlates study data for estimating the VE curve based on the log10-transformed fold rise in gpELISA ($S(1)$), using baseline gpELISA measure quartiles (W) as the BIP. A correlation of -0.50 was found between $S(1)$ and W among vaccine recipients. The data fit naturally into our multiphase sampling framework, with Y (herpes zoster infection) and Z (vaccination status) measured in the first phase, W available from all subjects in the immunogenicity set, and

$S(1)$ available among vaccinees in the immunogenicity set. Assuming a probit risk model $P(Y=1|Z) = \Phi(\beta_0 + \beta_1 Z + \beta_2 S(1) + \beta_3 Z)$, the proposed TPS estimator leads to estimated risk coefficients and their 95% Wald confidence intervals of $\hat{\beta}_0 = -2.65$ (95% CI= (-2.85, -2.45)), $\hat{\beta}_1 = 0.0096$ (95% CI= (-0.256, -0.275)), $\hat{\beta}_2 = -0.534$ (95% CI= (0.204, 0.863)), and $\hat{\beta}_3 = -1.233$ (95% CI= (-1.864, -0.602)). We show in Figure 1 the corresponding VE curve and 95% Wald confidence intervals assuming approximate normality of $\log\{1 - V E(s)\}$. $VE(s)$ increases from a VE value of -0.029 (95% CI= (-1.26, 0.53)) at a fold rise of 1 to a VE value of 0.9999 (95% CI= (0.9768, $1 - 6e^{-7}$)) at a maximum fold rise of 221.5, confirming the surrogate value of this biomarker. Vaccine efficacy curves generated based on other estimators TEL, WPS, and WEL are very similar to that generated from TPS. We show in Supplementary Table 10 the $VE(s)$ estimators and their standard error estimates for various s values for each estimator.

5. Discussion

In this paper we studied methodology for evaluating principal surrogates in vaccine efficacy trials through estimation of the vaccine efficacy curve, when baseline predictors used for predicting missing potential surrogate values are subsampled. The estimated vaccine efficacy curve can provide valuable information regarding the potential of developing efficacious vaccine through improving immunogenicity, and can also serve as an important tool in bridging or predicting the vaccine's effect in further trial settings (Gilbert and Huang, 2016). Having baseline predictors with good correlation with vaccine-induced immune responses is important for precision in surrogacy estimation. However, common baseline covariates that are easily available do not correlate well with vaccine-induced immune responses, making it critical to explore alternative baseline predictors such as baseline immune responses to the same antigens evaluated in the trial. New statistical methods development is needed to efficiently address the subsampling of these costly lab assay-based measures.

We phrase the problem under a multiphase sampling framework, and developed and compared several pseudo-score or estimated likelihood type estimators based on either full conditional likelihood or weighted likelihood. In general, we found that the pseudo-score type estimator performs similarly to the corresponding estimated likelihood estimator in a standard trial design with a BIP component, but can have much better efficiency in an augmented trial design with an additional CPV component. The same pattern has been observed earlier in a two-phase sampling setting in which vaccine-induced immune responses, but not the BIPs, were subsampled. Overall, both the pseudo-score and the estimated likelihood estimators based on full conditional likelihood (TPS and TEL) appear to perform optimally in a BIP-only design, while the pseudo-score estimator based on full conditional likelihood (TPS) performs best in a BIP+CPV design. Importantly, our estimators allow calculation of analytical variances valuable for inference and for guiding the design of immune surrogate studies, which has not previously been achieved.

The semiparametric estimators we studied estimate the distribution of W given X and the distribution of S given W and X in the population or among samples nonparametrically based on discrete W and X . This approach has the advantage of robustness to model assumption regarding the distribution of W and S . This is especially important for estimating

the distribution of S , W among samples (i.e. $S|X$, W , $\delta_2 = 1$ and $W|X$, $\delta_1 = 1$) as required in pseudo-score methods because these quantities vary with sampling schemes of S and W and are therefore hard to model with a parametric form. There is a trade-off of using this nonparametric procedure, however, in that its performance requires that observations of W and S not too sparse within the X or $\{X, W\}$ strata. Moreover, for the method to work, there must exist samples of S in each category of X and W and samples of W in each category of X . Based on various numerical explorations we had, for the nonparametric estimators of the conditional distribution of S and W to work well, one should have a few possible values for W and X , e.g. below 10. If there are large numbers of unique values for W and X , nonparametric smoothing techniques can be considered; for estimated likelihood methods, parametric (Gilbert and Hudgens, 2008) or semiparametric (Huang and Gilbert, 2011) estimators of the conditional distribution of W and S can also be entertained.

Under the multiphase sampling framework, another important question is the choice of sampling probabilities for W and S to optimize efficiency of the surrogacy evaluation under limited resources. While a full development of this issue is beyond the scope of this paper, we made a few interesting observations based on numerical explorations presented in the supplementary material. For example, given the interest in the biomarker's surrogacy, it seems natural to sample $S(1)$ with certainty when W is sampled if the costs for measuring W and S are comparable. This appears to be a sampling scheme with good performance in general (Supplementary Table 11). Having some CPV samples of $S(1)$ is also good for efficiency. After the measurement of a set of W and S , suppose additional resources become available for measurement, one can numerically assess the relative benefit of measuring extra S and/or W among vaccine or placebo recipients, as presented in Supplementary Table 12. Moreover, in simple linear regression of an outcome over a predictor, sampling extreme predictor values is known to be more efficient than simple random sampling. In our problem setting, we also observed potential advantages of oversampling extreme W and S on efficiency improvement (Supplementary Table 13). For designing a future surrogate marker study in practice, we recommend the use of numerical studies to evaluate various sampling schemes to be always considered.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by NIH grants R01 GM106177-01 and 2R37AI05465-10. We thank Dr. Peter Gilbert for helpful comments and suggestions.

References

- Burzykowski, T., Molenberghs, G., Buyse, M. The evaluation of surrogate endpoints. Springer Science & Business Media; 2006.
- Buyse M, Molenberghs G, Burzykowski T, et al. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics*. 2000; 1(1):49. [PubMed: 12933525]
- Chatterjee N, Chen Y, Breslow N. A pseudoscore estimator for regression problems with two-phase sampling. *JASA*. 2003; 98(461):158–168.

- Daniels M, Hughes M. Meta-analysis for the evaluation of potential surrogate markers. *Statistics in Medicine*. 1997; 16(17):1965–1982. [PubMed: 9304767]
- Follmann D. Augmented designs to assess immune response in vaccine trials. *Biometrics*. 2006; 62(4): 1161–1169. [PubMed: 17156291]
- Frangakis C, Rubin D. Principal stratification in causal inference. *Biometrics*. 2002; 58(1):21–29. [PubMed: 11890317]
- Freedman L, Graubard B, Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. *Statistics in Medicine*. 1992; 11(2):167–178. [PubMed: 1579756]
- Gabriel EE, Gilbert PB. Evaluating principal surrogate endpoints with time-to-event data accounting for time-varying treatment efficacy. *Biostatistics*. 2014; 15(2):251–265. [PubMed: 24337534]
- Gilbert P, Hudgens M. Evaluating candidate principal surrogate endpoints. *Biometrics*. 2008; 64(4): 1146–1154. [PubMed: 18363776]
- Gilbert PB, Gabriel EE, Huang Y, Chan IS. Surrogate endpoint evaluation: Principal stratification criteria and the prentice definition. *Journal of causal inference*. 2015; 3(2):157–175. [PubMed: 26722639]
- Gilbert PB, Gabriel EE, Miao X, et al. Fold rise in antibody titers by measured by glycoprotein-based enzyme-linked immunosorbent assay is an excellent correlate of protection for a herpes zoster vaccine, demonstrated via the vaccine efficacy curve. *Journal of Infectious Diseases*. 2014; 210(10):1573–1581. [PubMed: 24823623]
- Gilbert PB, Huang Y. Predicting overall vaccine efficacy in a new setting by re-calibrating baseline covariate and intermediate response endpoint effect modifiers of type-specific vaccine efficacy. *Epidemiologic Methods*. 2016; 5(1):93–112. [PubMed: 28154793]
- Horvitz DG, Thompson DJ. A generalization of sampling without replacement from a finite universe. *JASA*. 1952; 47(260):663–685.
- Huang Y, Gilbert PB. Comparing biomarkers as principal surrogate endpoints. *Biometrics*. 2011; 67(4):1442–1451. [PubMed: 21517791]
- Huang Y, Gilbert PB, Wolfson J. Design and estimation for evaluating principal surrogate markers in vaccine trials. *Biometrics*. 2013; 69(2):301–309. [PubMed: 23409839]
- Joffe M, Greene T. Related causal frameworks for surrogate outcomes. *Biometrics*. 2009; 65(2):530–538. [PubMed: 18759836]
- Li Y, Taylor J, Elliott M, Sargent D. Causal assessment of surrogacy in a meta-analysis of colorectal cancer trials. *Biostatistics*. 2011; 12(3):478. [PubMed: 21252079]
- Lin D, Fleming T, De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. *Statistics in medicine*. 1997; 16(13):1515–1527. [PubMed: 9249922]
- Pepe M, Fleming T. A nonparametric method for dealing with mismeasured covariate data. *Journal of the American Statistical Association*. 1991:108–113.
- Plotkin S. Correlates of protection induced by vaccination. *Clinical and Vaccine Immunology*. 2010; 17(7):1055. [PubMed: 20463105]
- Robins J, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992:143–155. [PubMed: 1576220]
- Robins JM, Rotnitzky A, Zhao LP. Estimation of regression coefficients when some regressors are not always observed. *JASA*. 1994; 89(427):846–866.
- Schmader KE, Levin MJ, Gnann JW, et al. Efficacy, safety, and tolerability of herpes zoster vaccine in persons aged 50–59 years. *Clinical infectious diseases*. 2012; 54(7):922–928. [PubMed: 22291101]
- Wolfson J, Gilbert P. Statistical identifiability and the surrogate endpoint problem, with application to vaccine trials. *Biometrics*. 2010; 66(4):1153–1161. [PubMed: 20105158]

Figure 1. TPS estimator of VE and 95% Wald confidence interval as a function of the vaccine-induced fold rise of gpELISA antibody titers from baseline to week 6.

Table 1

Finite sample performance of various estimators in BIP-only designs

γ_{w_v}	γ_{s_v}	γ_{w_p}	γ_{s_p}	Method	β_0	β_1	β_2	β_3	$VE(s_{0.1})^*$	$VE(s_{0.3})^*$	$VE(s_{0.5})^*$	$VE(s_{0.7})^*$	$VE(s_{0.9})^*$
Bias $\times 100$													
1	1	1	0	TPS	0.35	0.76	-0.29	-0.17	-2.91	-0.76	-0.45	-0.82	-2.21
				TEL	-0.78	2.00	0.07	-0.58	-3.45	-0.95	-0.47	-0.75	-2.11
				WPS	0.26	1.31	-0.34	-0.28	-4.56	-1.14	-0.64	-1.27	-3.80
				WEL	-0.17	1.74	-0.19	-0.43	-4.82	-1.20	-0.64	-1.27	-4.34
2	2	2	0	TPS	0.17	0.67	-0.22	-0.17	-2.36	-0.61	-0.35	-0.64	-1.69
				TEL	-0.24	0.93	-0.07	-0.27	-2.31	-0.61	-0.33	-0.58	-1.55
				WPS	-0.05	0.89	-0.15	-0.24	-2.83	-0.70	-0.37	-0.70	-1.95
				WEL	-0.16	0.99	-0.12	-0.28	-2.86	-0.71	-0.37	-0.69	-1.94
5	5	5	0	TPS	-0.16	0.54	-0.09	-0.17	-1.84	-0.47	-0.26	-0.48	-1.30
				TEL	-0.24	0.55	-0.05	-0.18	-1.77	-0.44	-0.24	-0.46	-1.25
				WPS	-0.19	0.50	-0.07	-0.17	-1.84	-0.45	-0.24	-0.47	-1.32
				WEL	-0.22	0.53	-0.06	-0.17	-1.85	-0.46	-0.24	-0.47	-1.31
10	10	10	0	TPS	-0.30	0.16	-0.03	-0.06	-1.42	-0.35	-0.23	-0.47	-1.23
				TEL	-0.31	0.17	-0.03	-0.06	-1.41	-0.34	-0.22	-0.47	-1.22
				WPS	-0.30	0.14	-0.03	-0.05	-1.41	-0.34	-0.22	-0.47	-1.24
				WEL	-0.31	0.15	-0.02	-0.05	-1.41	-0.34	-0.22	-0.47	-1.23
var $\times 100$													
1	1	1	0	TPS	13.22	13.03	1.56	1.58	3.04	0.50	0.29	0.60	1.94
				TEL	11.59	13.34	1.39	1.63	3.20	0.51	0.30	0.65	1.94
				WPS	16.33	20.67	1.97	2.52	5.14	0.65	0.32	0.94	4.26
				WEL	16.72	20.85	2.03	2.54	5.75	0.67	0.34	1.15	13.88
2	2	2	0	TPS	9.83	10.69	1.17	1.29	2.45	0.45	0.28	0.50	1.12
				TEL	8.47	10.09	1.01	1.22	2.33	0.45	0.28	0.50	1.10
				WPS	10.19	12.83	1.21	1.54	3.02	0.50	0.29	0.56	1.33
				WEL	10.23	12.79	1.22	1.54	3.01	0.50	0.29	0.56	1.35
5	5	5	0	TPS	6.83	8.59	0.81	1.03	1.97	0.42	0.27	0.43	0.86
				TEL	6.35	8.27	0.75	0.99	1.91	0.42	0.27	0.43	0.84

γ_{w_y}	γ_{s_y}	γ_{w_p}	γ_{s_p}	Method	β_0	β_1	β_2	β_3	$VE(s_{0,1})^*$	$VE(s_{0,3})^*$	$VE(s_{0,5})^*$	$VE(s_{0,7})^*$	$VE(s_{0,9})^*$
10	10	10	0	WPS	6.67	8.79	0.79	1.05	2.04	0.43	0.27	0.44	0.87
				WEL	6.67	8.77	0.79	1.05	2.04	0.43	0.27	0.44	0.88
				TPS	5.62	7.48	0.67	0.90	1.71	0.40	0.27	0.41	0.76
				TEL	5.56	7.44	0.66	0.89	1.71	0.40	0.27	0.41	0.75
1	1	1	0	WPS	5.61	7.49	0.67	0.90	1.73	0.40	0.27	0.41	0.76
				WEL	5.61	7.48	0.67	0.90	1.73	0.40	0.27	0.41	0.76
				coverage of 95% CI									
				TPS	96.7	96.3	96.7	96.3	96.2	95.5	95.6	96.0	95.8
2	2	2	0	TEL	96.8	96.8	96.5	96.2	96.5	95.7	95.4	95.9	96.0
				WPS	96.2	96.0	96.3	96.1	95.5	95.8	96.0	95.9	96.2
				WEL	96.8	96.2	96.8	96.3	95.8	96.0	95.8	96.1	96.3
				TPS	96.2	96.1	96.0	95.9	95.9	95.7	95.3	96.2	96.2
5	5	5	0	TEL	96.2	96.1	96.3	95.9	96.2	96.0	95.2	96.0	95.9
				WPS	95.9	95.8	96.1	96.0	95.7	95.8	95.4	96.0	96.4
				WEL	96.3	96.0	96.2	96.0	95.7	95.8	95.3	96.0	96.4
				TPS	96.1	95.8	96.1	95.8	95.6	95.6	95.2	95.9	95.7
10	10	10	0	TEL	96.1	95.7	96.1	95.8	95.8	95.5	95.2	95.7	95.7
				WPS	96.2	95.7	96.3	95.7	95.9	95.5	95.4	95.8	95.9
				WEL	96.2	95.7	96.3	95.7	96.0	95.4	95.3	95.8	95.9
				TPS	96.2	95.8	95.8	95.5	95.9	95.5	95.3	95.6	95.8
10	10	10	0	TEL	96.2	95.9	95.9	95.7	96.0	95.5	95.2	95.4	95.7
				WPS	96.2	95.9	96.0	95.5	96.1	95.6	95.3	95.5	95.8
				WEL	96.2	96.0	96.0	95.6	96.2	95.5	95.2	95.5	95.7
				TPS	96.2	96.0	96.0	95.6	96.2	95.5	95.2	95.5	95.7

$VE(s_{\alpha})^*$ is $VE(s)$ for s the $\alpha \times 100\%$ percentile of $S(1)$

Table 2

Finite sample performance of various estimators in BIP+CPV designs

γ_{w_v}	γ_{s_y}	γ_{w_p}	γ_{s_p}	Method	β_0	β_1	β_2	β_3	$VE(s_{0,1})^*$	$VE(s_{0,3})^*$	$VE(s_{0,5})^*$	$VE(s_{0,7})^*$	$VE(s_{0,9})^*$
Bias $\times 100$													
1	1	1	1	TPS	0.03	0.84	-0.19	-0.21	-2.54	-0.68	-0.39	-0.69	-1.81
				TEL	-1.07	2.29	0.14	-0.65	-3.68	-1.06	-0.54	-0.82	-2.48
				WPS	2.27	-0.66	-1.39	0.76	-11.31	-2.30	-1.73	-4.33	-20.99
				WEL	-0.43	2.05	-0.82	0.18	-23.00	-4.27	-2.73	-6.38	-35.54
2	2	2	2	TPS	-0.14	0.70	-0.12	-0.20	-2.10	-0.55	-0.31	-0.56	-1.46
				TEL	-0.64	1.32	0.04	-0.39	-2.55	-0.71	-0.36	-0.58	-1.53
				WPS	1.10	-0.27	-0.67	0.27	-4.09	-0.97	-0.70	-1.54	-4.49
				WEL	0.08	0.76	-0.51	0.12	-8.52	-1.81	-1.21	-2.54	-10.02
5	5	5	5	TPS	-0.43	0.66	0.01	-0.22	-1.70	-0.44	-0.23	-0.42	-1.14
				TEL	-0.66	0.97	0.08	-0.31	-1.94	-0.52	-0.26	-0.43	-1.17
				WPS	-0.23	0.55	-0.07	-0.17	-2.01	-0.50	-0.27	-0.53	-1.47
				WEL	-0.59	0.90	-0.01	-0.23	-2.84	-0.73	-0.41	-0.76	-2.03
10	10	10	10	TPS	-0.51	0.32	0.04	-0.12	-1.30	-0.33	-0.21	-0.41	-1.08
				TEL	-0.66	0.52	0.09	-0.18	-1.48	-0.38	-0.22	-0.43	-1.13
				WPS	-0.47	0.30	0.03	-0.11	-1.33	-0.33	-0.21	-0.42	-1.12
				WEL	-0.63	0.46	0.07	-0.15	-1.61	-0.40	-0.25	-0.49	-1.30
var $\times 100$													
1	1	1	1	TPS	9.32	11.32	1.11	1.37	2.62	0.47	0.29	0.53	1.24
				TEL	10.78	13.79	1.31	1.69	3.36	0.52	0.32	0.81	6.89
				WPS	32.92	45.82	4.10	5.61	120.26	2.02	0.68	6.59	587.96
				WEL	48.14	62.53	5.91	7.59	357.15	4.11	1.02	11.59	1082.88
2	2	2	2	TPS	7.33	9.48	0.87	1.14	2.20	0.44	0.28	0.47	0.95
				TEL	7.64	10.33	0.92	1.25	2.45	0.46	0.28	0.50	1.09
				WPS	15.15	21.39	1.85	2.59	6.07	0.69	0.33	1.04	6.19
				WEL	22.90	30.30	2.81	3.68	108.92	1.49	0.49	3.18	215.16
5	5	5	5	TPS	5.45	7.74	0.65	0.93	1.82	0.41	0.27	0.41	0.76
				TEL	5.66	8.22	0.68	0.99	1.94	0.42	0.27	0.43	0.81

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

γ_{w_1}	γ_{s_1}	γ_{w_2}	γ_{s_2}	Method	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$VE(s_{0,1})^*$	$VE(s_{0,3})^*$	$VE(s_{0,5})^*$	$VE(s_{0,7})^*$	$VE(s_{0,9})^*$
10	10	10	10	WPS	6.66	9.67	0.80	1.16	2.26	0.45	0.28	0.46	0.96
				WEL	9.33	12.76	1.13	1.54	3.03	0.50	0.29	0.56	1.38
				TPS	4.56	6.61	0.55	0.80	1.55	0.38	0.27	0.39	0.69
				TEL	4.92	7.09	0.59	0.86	1.64	0.39	0.27	0.41	0.75
				WPS	4.70	6.83	0.56	0.83	1.59	0.39	0.27	0.40	0.72
				WEL	5.65	7.91	0.69	0.96	1.80	0.40	0.28	0.44	0.83
1	1	1	1	coverage of 95% CI									
				TPS	96.5	96.3	96.2	95.8	96.1	95.6	95.3	95.9	95.9
				TEL	96.5	96.4	96.3	96.1	96.3	95.9	95.4	95.6	96.0
				WPS	95.1	95.1	95.4	95.2	94.1	95.9	96.8	95.2	95.0
				WEL	95.1	95.0	95.1	95.2	93.3	95.8	97.4	94.2	93.7
				TPS	96.0	95.9	95.8	95.7	95.9	95.8	95.3	96.0	96.1
2	2	2	2	TEL	95.9	95.8	95.8	95.8	95.9	95.9	95.2	95.8	95.7
				WPS	95.4	95.0	95.8	95.5	95.1	95.8	95.6	96.1	95.9
				WEL	95.2	95.3	95.7	95.6	94.8	96.1	96.2	95.7	95.0
				TPS	96.0	95.7	95.9	95.5	95.8	95.6	95.3	95.7	95.6
				TEL	96.0	95.5	96.0	95.6	95.8	95.9	95.3	95.8	96.0
				WPS	96.2	95.8	96.2	95.8	95.7	95.5	95.4	95.9	96.2
5	5	5	5	WEL	95.7	95.7	95.7	95.6	95.7	95.9	95.5	95.9	95.6
				TPS	95.8	95.4	95.2	95.2	95.6	95.8	95.2	95.5	95.6
				TEL	95.6	95.5	95.2	95.3	95.9	95.8	95.1	95.1	95.4
				WPS	95.7	95.7	95.5	95.4	95.8	95.6	95.2	95.5	95.7
				WEL	95.7	95.6	95.7	95.2	96.0	95.8	95.1	95.3	95.5
				TPS	95.7	95.6	95.7	95.2	96.0	95.8	95.1	95.3	95.5

$VE(s_{\alpha})^*$ is $VE(s)$ for the $\alpha \times 100\%$ percentile of $S(1)$

Efficiency TPS relative to estimator A = variance of A/variance of TPS.

	β_1			β_2			β_3			β_4			χ_{d_1}	χ_{d_2}	χ_{s_1}	χ_{s_2}
	WEL	WPS	TEL	WEL	WPS	TEL	WEL	WPS	TEL	WEL	WPS	TEL				
1	2.34	1.23	1.80	2.29	1.04	1.69	2.33	1.23	1.79	2.25	1.04	1.66	10	10	2	2
5	2.51	2.10	1.03	2.76	2.25	1.04	2.48	2.08	1.03	2.69	2.20	1.03	2	2	5	5
10	1.66	1.25	1.07	1.74	1.23	1.04	1.65	1.25	1.06	1.71	1.22	1.04	5	5	5	5
15	1.46	1.07	1.23	1.50	1.03	1.23	1.46	1.07	1.23	1.48	1.03	1.22	10	10	5	5
20	1.20	1.03	1.08	1.25	1.03	1.08	1.20	1.03	1.07	1.24	1.03	1.08	10	10	10	10

Table 4

Efficiency of TPS relative to various alternative estimators for estimating $VE(S(1))$.

γ_{w_v}	γ_{s_v}	γ_{w_p}	γ_p	$VE(s_{0,1})^*$				$VE(s_{0,3})^*$				$VE(s_{0,5})^*$				$VE(s_{0,7})^*$				$VE(s_{0,9})^*$			
				TPS	TEL	WPS	WEL	TPS	TEL	WPS	WEL	TPS	TEL	WPS	WEL	TPS	TEL	WPS	WEL	TPS	TEL	WPS	WEL
BIP-only Design																							
1	1	1	0	1.00	1.05	1.69	1.89	1.00	1.02	1.31	1.36	1.00	1.03	1.11	1.15	1.00	1.07	1.56	1.90	1.00	1.00	2.19	7.14
1	1	5	0	1.00	1.04	1.14	1.14	1.00	1.02	1.08	1.09	1.00	1.01	1.01	1.01	1.00	1.03	1.05	1.05	1.00	1.02	1.02	1.03
1	1	10	0	1.00	1.06	1.16	1.16	1.00	1.02	1.07	1.07	1.00	1.01	1.01	1.01	1.00	1.03	1.09	1.08	1.00	1.05	1.14	1.35
5	1	5	0	1.00	1.05	1.04	1.11	1.00	1.02	1.03	1.06	1.00	1.01	1.00	1.02	1.00	1.03	1.00	1.04	1.00	1.03	0.99	1.05
10	1	10	0	1.00	1.06	1.00	1.06	1.00	1.02	1.00	1.02	1.00	1.02	1.00	1.02	1.00	1.09	1.00	1.08	1.00	1.26	1.00	1.22
2	2	2	0	1.00	0.95	1.23	1.23	1.00	0.98	1.11	1.11	1.00	1.00	1.01	1.02	1.00	0.99	1.11	1.11	1.00	0.98	1.18	1.20
2	2	5	0	1.00	0.97	1.04	1.04	1.00	0.99	1.04	1.04	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.99	1.00	1.01
2	2	10	0	1.00	1.01	1.06	1.05	1.00	1.00	1.02	1.03	1.00	1.00	1.00	1.00	1.00	1.01	1.02	1.01	1.00	1.01	1.04	1.02
5	5	2	0	1.00	0.98	1.29	1.29	1.00	0.99	1.11	1.11	1.00	1.01	1.02	1.03	1.00	1.01	1.17	1.18	1.00	1.01	1.37	1.42
5	5	5	0	1.00	0.97	1.04	1.04	1.00	0.99	1.02	1.02	1.00	1.00	1.00	1.00	1.00	0.98	1.01	1.01	1.00	0.97	1.01	1.01
5	5	10	0	1.00	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
10	10	10	0	1.00	1.00	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00
BIP+CPV Design																							
1	1	1	1	1.00	1.28	45.86	136.21	1.00	1.11	4.31	8.76	1.00	1.11	2.36	3.58	1.00	1.53	12.36	21.74	1.00	5.55	473.59	872.24
1	1	5	5	1.00	2.19	2.08	13.60	1.00	1.46	1.47	3.63	1.00	1.13	1.06	1.70	1.00	1.61	1.46	4.89	1.00	2.24	2.06	33.91
1	1	10	10	1.00	4.00	1.72	6.82	1.00	2.04	1.32	2.82	1.00	1.75	1.03	2.05	1.00	8.32	1.24	11.64	1.00	454.01	1.47	715.07
5	1	5	1	1.00	1.20	1.12	1.61	1.00	1.08	1.05	1.25	1.00	1.04	1.01	1.10	1.00	1.14	1.05	1.44	1.00	1.22	1.10	2.27
10	1	10	1	1.00	1.20	1.01	1.28	1.00	1.08	1.00	1.11	1.00	1.04	1.00	1.05	1.00	1.19	1.01	1.25	1.00	1.57	1.01	1.83
2	2	2	2	1.00	1.11	2.76	49.45	1.00	1.04	1.58	3.42	1.00	1.02	1.18	1.74	1.00	1.08	2.23	6.82	1.00	1.16	6.54	227.20
2	2	5	5	1.00	1.33	1.47	2.91	1.00	1.12	1.21	1.67	1.00	1.04	1.03	1.19	1.00	1.19	1.23	2.07	1.00	1.34	1.52	4.00
2	2	10	10	1.00	1.94	1.27	2.69	1.00	1.33	1.11	1.58	1.00	1.11	1.01	1.20	1.00	1.62	1.11	2.23	1.00	3.21	1.21	9.14
5	5	2	2	1.00	1.03	2.21	3.17	1.00	1.01	1.39	1.60	1.00	1.01	1.17	1.28	1.00	1.02	2.36	3.30	1.00	1.04	20.40	60.57
5	5	5	5	1.00	1.06	1.24	1.66	1.00	1.02	1.09	1.23	1.00	1.01	1.01	1.05	1.00	1.04	1.12	1.36	1.00	1.07	1.26	1.80
5	5	10	10	1.00	1.24	1.07	1.47	1.00	1.08	1.03	1.15	1.00	1.02	1.00	1.04	1.00	1.12	1.04	1.24	1.00	1.23	1.07	1.48
10	10	10	10	1.00	1.06	1.03	1.17	1.00	1.02	1.01	1.05	1.00	1.01	1.00	1.02	1.00	1.04	1.02	1.11	1.00	1.08	1.04	1.21

Efficiency of TPS relative to estimator A = variance of A/variance of TPS.

$(1 - S) \times 100\%$ the proportion of S is $(1 - S) \times 100\%$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript