Estimating Malaria Vaccine Efficacy in the Absence of a Gold Standard Case Definition: Mendelian Factorial Design

Raiden B. Hasegawa[*] and Dylan S. Small[†]

September 6, 2020

Abstract

Accurate estimates of malaria vaccine efficacy require a reliable definition of a malaria case. However, the symptoms of clinical malaria are unspecific, overlapping with other childhood illnesses. Additionally, children in endemic areas tolerate varying levels of parasitemia without symptoms. Together, this makes finding a gold-standard case definition challenging. We present a method to identify and estimate malaria vaccine efficacy that does not require an observable gold-standard case definition. Instead, we leverage genetic traits that are protective against malaria but not against other illnesses, e.g., the sickle cell trait, to identify vaccine efficacy in a randomized trial. Inspired by Mendelian randomization, we introduce Mendelian factorial design, a method that augments a randomized trial with genetic variation to produce a natural factorial experiment, which identifies vaccine efficacy under realistic assumptions. A robust, covariance adjusted estimation procedure is developed for estimating vaccine efficacy on the risk ratio and incidence rate ratio scales. Simulations suggest that our estimator has good performance whereas standard methods are systematically biased. We demonstrate that a combined estimator using both our proposed estimator and the standard approach yields significant improvements when the Mendelian factor is only weakly protective.

Keywords: Vaccine Efficacy; Causal Inference; Malaria; Mendelian Randomization; Sickle Cell; Factorial Design

---

[*]*Address for correspondence:* Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, (e-mail: `raiden@wharton.upenn.edu`)

[†]Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, (e-mail: `dsmall@wharton.upenn.edu`)

1 Introduction

In 2017, there were an estimated 219 million cases of malaria of which 92% occurred in Africa and an estimated 435,000 malaria related deaths of which 93% occurred in Africa; nearly every case of malaria in Africa was cause by the parasite *Plasmodium falciparum* (World Health Organization, 2018). Pregnant women and children under the age of 5 are the most vulnerable groups affected by malaria. Of more than 30 vaccines under development, the only vaccine to undergo a pivotal phase III trial is the pre-erythrocytic vaccine RTS,S which has shown to have limited efficacy ($39-50\%$ per-protocol reduction in incidence rates; Mahmoudi and Keshavarz (2017); The RTS,S Clinical Trials Partnership (2014)). Consequently, the continued development of an efficacious P. *falciparum* malaria vaccine has the potential for substantial public health impacts. With so many potential vaccines in the development pipeline, a critically important statistical challenge is to develop methods for estimating vaccine efficacy. Until now, accurate estimation of vaccine efficacy against clinical outcomes attributable to P. *falciparum* malaria has required defining reliable case definitions, a task that is made difficult by the unspecific presentation of malaria in endemic areas.

In the absence of a gold standard case definition, efficacy is usually assessed by choosing an inexact case definition which may falsely exclude cases attributable malaria and falsely include non-malaria cases. The established definition of clinical malaria in malaria prevention trials is the presence of a fever with a temperature $\geq 37.5^{\circ}$C and a P. *falciparum* parasite density above a certain threshold, e.g., 2500 or 5000 parasites per μl of blood (Ter Kuile et al., 2003; The RTS,S Clinical Trials Partnership, 2011; Olotu et al., 2013; Bejon et al., 2013). Case definitions of this form inevitably exclude some true cases and include some false cases due to heterogeneity in immunity and endemicity, fever killing effects, and parasite density measurement error. With specificity < 1, estimates of vaccine efficacy will be biased downward. It has been shown in simulations that such case classification errors have the potential to introduce substantial bias in many settings (Small et al., 2010). Real trial data suggests that these challenges with the standard case definition often go unaddressed. In a multi-site pooled analysis of phase II RTS,S trial data using a fixed 2500 parasite per μl cutoff across all study sites, Bejon et al. (2013) reports estimates of vaccine efficacy against malaria that were as high as 60% in sites with low parasite prevalence and as low as 4% in high

parasite prevalence sites. The authors suggest a biological reason for this pattern: RTS,S prevents fevers in only a portion of mosquito bites and in areas where parasite prevalence is high, children are likely to be bitten more often by infected mosquitos. However this pattern is also consistent with bias due to the fixed case definition having lower specificity in higher prevalence areas where, due to improved immunity, children can carry higher parasite loads without fever. In fact, pivotal phase III RTS,S trial data that used the higher specificity threshold of 5000 parasites per μl in the primary case definition found that the variation in vaccine efficacy could not be explained by variation in exposure to malaria (The RTS,S Clinical Trials Partnership, 2014). The standard case definition leaves much unanswered: is the heterogeneity in vaccine efficacy epidemiologically important or just an artifact of bias arising from an inexact case definition?

To overcome the challenges that accompany inexact case definitions, this paper introduces a new method for identifying malaria vaccine efficacy using natural genetic variation, which we call *Mendelian factorial design.* Importantly, the method does not depend on an inexact case definition based on the parasite density, instead using all fevers (or deaths) with any level of parasitemia to estimate efficacy. To identify vaccine efficacy, the new method requires finding and recording genetic variants that provide specific protection against clinical malaria and operate through a different biological pathway than the vaccine being evaluated. A running example in this paper is the sickle cell trait, a hemoglobinopathy that has been shown to protect against malaria at the blood-stage (erythrocytic) of the infection and the RTS,S vaccine, which confers protection against malaria at the pre-erythrocytic, liver-stage of the infection. Symptoms, such as fever, develop in the erythrocytic-stage of the disease. There are many other vaccine types, e.g., transmission blocking and erythrocytic vaccines, and genetic variants that might satisfy these identifying conditions (Ndila et al., 2018). We will discuss these occasionally throughout the paper and more thoroughly in the discussion.

In the following section, we introduce Mendelian factorial design informally in the context of the more familiar use of genetic variants as instrumental variables in *Mendelian randomization* studies, to which it has many parallels. In §3, we present a precise definition of vaccine efficacy and malaria-attributable fevers in a potential outcome framework and propose an identification strategy that holds under a few realistic assumptions. We then provide a simple $\sqrt{n}$-consistent and asymptotically normal covariate-adjusted estimator that is ro-

2

bust to model misspecification and assess its performance in a simulation study over a range of settings. We develop an improved "bounded" estimator that combines the strengths of our estimator with that of the (biased) standard estimator. We demonstrate that it provides nearly uniform improvement over both estimators. In Appendix A of the Supplemental Web Materials we extend the Mendelian factorial design to identify and estimate vaccine efficacy in time-to-first event studies.

2 Mendelian Factorial Design: Parallels with Mendelian Randomization

When an observed association between a non-randomized exposure and an outcome may be confounded by an unobserved common cause, attributing the association to a causal effect may be misleading (Rosenbaum, 2002). An instrumental variable (IV) is a covariate that is associated with the exposure but whose only association to the outcome is through a direct pathway, causal or not, to the exposure (Martens et al., 2006). The IV takes the place of the physical randomization in a randomized trial, influencing only the assignment of subjects to exposure or control, setting up a natural experiment and providing an avenue for estimating the causal effect of the non-randomized exposure on the outcome (Rassen et al., 2009). When genetic variants are used as instrumental variables, the corresponding collections of methods are often labeled as the *Mendelian randomization* (MR) approach (Smith and Ebrahim, 2008). For example, Kang et al. (2013) proposed using the hemoglobin S variant (HbS) as an instrument to identify the causal effect of malaria on stunted growth. It is known that heterozygote carriers (HbAS, sickle cell trait) receive protection against clinical malaria whilst those without the variant (HbAA) do not. The first column of Table 1 summarizes the assumptions required of HbAS status such that it can be used in a MR study to identify the causal effect of malaria on stunting. Assumption (1) requires that HbAS status in fact influences exposure; assumption (2) ensures that HbAS status can be treated "as-if" it were randomized; and assumption (3) says that HbAS only effects stunted growth through its influence on exposure. Assumption (4) is required for point identification of the causal effect, although there are other alternative assumptions that can be used to point identify a

causal effect such as monotonicity of the effect (Burgess et al., 2017).

Like MR, Mendelian factorial design (MFD) leverages genetic variation to identify a causal effect of interest – in this instance, the efficacy of a vaccine, or roughly, the proportion reduction of disease-attributable outcomes in a population when the vaccine is applied (Lachenbruch, 1998). However, instead of addressing the bias that often accompanies a non-randomized exposure, MFD attends to the bias in estimating treatment efficacy arising from an inexact case definition. The natural experiment arising from MR is used for causal inference in the absence of an actual randomized experiment, enabling the investigator to distinguish causal effects from unobserved confounding. The same natural experiment is employed in MFD to augment a two-arm randomized trial and create a simple 2×2 factorial experiment. The factorial design allows the investigator to distinguish efficacy against disease-attributable outcomes even when the case definition incorrectly classifies a material number of non-disease outcomes as cases.

MFD is distinct from factorial MR, which was introduced in the original extended exposition of MR (Davey Smith and Ebrahim, 2003) as a possible way to strengthen a genetic instrument by looking at interacting independent Mendelian genes. More recent work has investigated using factorial MR as a potential way to estimate interaction effects of multiple exposures (Rees et al., 2019; North et al., 2019).

Although MR and MFD address different sources of bias in different settings, their designs share many parallel features. This is illustrated in Table 1, where each assumption of MR in the malaria stunting study corresponds to a closely related assumption that underlies a MFD study of vaccine efficacy against malaria-attributable fever. In more general terms, Assumption (1) says that the *Mendelian gene* or *Mendelian factor*, e.g., HbAS, is relevant. In the MR study this means that it is associated with the non-randomized exposure and in the MFD study this implies that the Mendelian gene is protective against the disease-attributable outcome. Assumption (2) requires that the Mendelian gene be unconfounded with the outcome of interest or be "as-if" randomized, which implies the former. For the MR study, assumption (3) says that the Mendelian gene has no pleiotropic effects (Smith and Ebrahim, 2008), that is, HbAS only effects stunted growth through its influence on malaria and not through its influence on another modifiable exposure or stunting itself. The parallel assumption for a MFD study is that the Mendelian factor protects against outcomes of any-

Table 1: Parallel assumptions for Mendelian randomization and Mendelian factorial design.

Assumption	*Mendelian Randomization*	**Mendelian Factorial Design**
(1)	• HbAS associated with malaria	• HbAS has protective efficacy against malaria-attributable fever
(2)	• No unmeasured confounders associated with HbAS and stunted growth	• HbAS "as-if" randomized
(3)	• Only direct pathway from HbAS to stunted growth is through malaria	• Protection provided by HbAS is specific to malaria-attributable fever
(4)	• HbAS does not modify the effect of malaria on stunting	• No interaction effect between HbAS and vaccine (i.e., independent protective pathways)

cause only through reducing disease-attributable outcomes. Finally, assumption (4) says that the causal effect in the MR study does not vary over different levels of the Mendelian gene, or in other words, they do not interact. Similarly, the corresponding assumption in a MFD study is that the vaccine and Mendelian factor do not interact. In other words, the vaccine prevents the same proportion of disease-attributable outcomes at different levels of the Mendelian gene. For example, this assumption is plausible when a malaria vaccine and HbAS provide protection against malaria-attributable fevers through independent biological pathways.
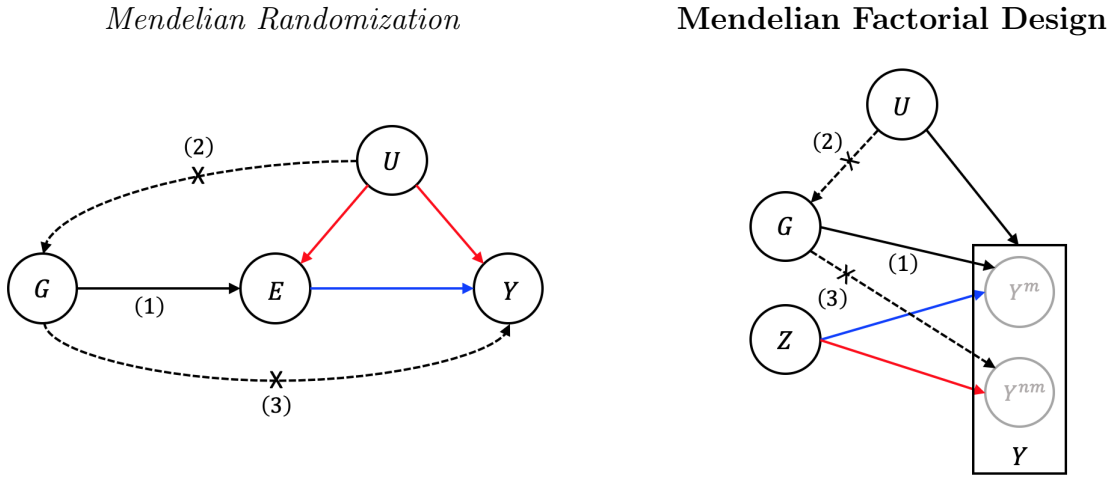

Figure 1: Causal diagrams for MR (left) and MFD (right). Blue arrows indicate the causal quantities of interest and red arrows indicate confounding addressed by each design. Gray variables Y^m and Y^{nm} are unobservable but Y, which doesn't distinguish between them is. Assumptions (1), (2) and (3) correspond to the similarly numbered assumptions in Table 1.

.

Assumptions (1)-(3) in Table 1 may be better understood encoded in a causal diagram. The causal diagram for MR and MFD are in the left and right panels of Figure 1, respectively. G is a Mendelian gene (or factor), U are unobserved confounders, Z is a randomized vaccine, and E is a non-randomized exposure. Y is the outcome of interest – in the MR study it is stunted growth and in the MFD study it is clinical malaria. Y^m and Y^{nm} are malaria-attributable and non-malaria fevers, respectively. We define these more precisely in §3. Finally, the blue arrows indicate the causal quantities of interest and the red arrows indicate

the confounding addressed by each design. The causal diagram for MFD is a little bit unusual. First, unlike the MR causal diagram, naive estimation of treatment efficacy will be biased not when the magnitude of the red arrow is non-zero but when it is not equal to the magnitude of the blue arrow. Second, Y^m and Y^{nm} are grayed out, emphasizing that we don't observed them, but instead only observe the outcome that is not distinguished by cause, Y. This is represented by the black bounding box. The MFD diagram hints at how the factorial structure and the absence of an arrow between G and Y^{nm} may be used to identify the arrow between Z and Y^m, the vaccine efficacy.

3 A Robust Framework for Estimating Vaccine Efficacy: Risk Ratios and Incidence Rate Ratios

3.1 Notation: Observed Data

Let $j = 1, \ldots, J$ indicate sites in a multi-site randomized control trial (RCT) and $i = 1, \ldots, I_j$ indicate subjects at each center; then, ij uniquely identifies each subject in the study. We let the total number of subjects in the trial be $n = \sum_j I_j$. For subject ij, let $S_{ij} = j$ be the site in which the subject was observed; let $Y_{ij} \in \{0, 1\}$ be an observed fever or death with any parasitemia; let $Z_{ij} \in \{0, 1\}$ indicate treatment/vaccine status and $G_{ij} \in \{0, 1\}$ indicate sickle cell variant status ($G_{ij} = 1$ if HbAS, $G_{ij} = 0$ if HbAA); and let $X_{ij} \in \mathbb{R}^d$ be a d-dimensional vector of baseline characteristics. Because G is assigned at conception, careful consideration of what variables constitute "baseline" variables should be made. Let D_{ij} and U_{ij} indicate the malaria parasite density in the blood and the level of non-malaria infectious agents measured at time of follow up, respectively. Define the observed data vector $O_{ij} = (X_{ij}, Z_{ij}, G_{ij}, Y_{ij})$. While U_{ij} is generally not observed in malaria trials, D_{ij} is. In Section 3.2.3, we describe the setting and notation when considering a study with multiple follow-ups. However, because the methods developed in this paper do not require or explicitly model parasite density we omit it from O_{ij}. We drop the subscripts and write $O = (X, Z, G, Y)$ to denote a random draw from a specified population. We will use boldface to indicate corresponding vector and matrix quantities that collect the data of subjects over each site or the entire trial. For example, $\mathbf{Y}_j$ is the $I_j \times 1$ vector that collects all the outcome

data for subjects at site j and $\mathbf{Y} = (\mathbf{Y}_1^T, \ldots, \mathbf{Y}_J^T)^T$ is the $n \times 1$ vector that collects the outcome data for all subjects in the multi-site trial. For an example of a matrix quantity, $\mathbf{X}_j$ is an $I_j \times d$ matrix and $\mathbf{X}$ is a $n \times d$ matrix.

Multi-site RCTs, i.e., block randomized, are a common design for vaccine efficacy trials, such as the Phase III RTS,S trial (The RTS,S Clinical Trials Partnership, 2011). Another design that has commonly been employed in studying the protective efficacy of interventions such as insecticide treated bed nets is the clustered RCT (Ter Kuile et al., 2003). The notation is easily adapted for this design, letting j indicate a cluster and $Z_{ij} = Z_j$ for all i, j.

3.2 Potential Outcomes and Malaria-Attributable Fever (or Death)

3.2.1 Pitfalls of Standard Case Definitions

The standard case definition of clinical malaria is the presence of a fever ($Y = 1$) and a parasite density above some threshold d ($D > d$). The WHO recommends setting d high enough to achieve specificity $> 80\%$ at all sites to avoid severe underestimation of vaccine efficacy (Moorthy et al., 2007). Sensitivity is also considered when determining d to avoid under-powered studies. The prevailing method used to determine these thresholds is to model risk of fever as a continuous function of observed parasite density among community controls and clinically suspected cases (Smith et al., 1994). However, computing sensitivity and specificity of a case definition requires an estimate of the malaria attributable fraction of fevers (MAFF). Lee and Small (2018) show that obtaining unbiased estimates of MAFF in the presence of measurement error of D and fever killing effects on parasite density is a difficult task, especially when malaria and non-malaria infections can work in conjunction to trigger a fever. Still, when unbiased estimates of MAFF can be obtained, this case definition will, by design, result in false positive and false negative cases, which may bias corresponding estimates of vaccine efficacy. Heterogeneous immunity and pyrogenic thresholds further complicate estimating treatment efficacy using case definitions that depend on a fixed threshold for D. In practice, standard thresholds without, it seems, careful estimation of specificity and sensitivity such as 2500 parasites per μl (Olotu et al., 2013) and 5000 parasites per μl are commonly used (The RTS,S Clinical Trials Partnership, 2011). As far as we are aware, there are no sufficiently specific case-definitions for death attributable to malaria, the most

severe effect of a malaria infection (Moorthy et al., 2007).

3.2.2 Defining Cases Using Potential Outcomes

Potential outcomes are a useful framework to precisely define causal effects of interest (Rubin, 2005). The potential outcomes that we define now are closely related to those defined in Lee and Small (2018). In what follows, we consider Y to indicate the presence of fever but note that the framework we describe is also suitable when Y indicates death. The aforementioned authors motivate their framework with a biological model of malaria and the notion of a pyrogenic threshold (Gravenor and Kwiatkowski, 1998). A pyrogenic threshold can roughly be defined as a level of malaria parasite density above which a fever will be produced. Below the threshold, the infection will not be strong enough to trigger a fever. This threshold may vary from individual to individual based on heterogeneous immunity to symptomatic malaria. In general, we can think of D and U as having thresholds above which a malaria or non-malaria infection, respectively, is strong enough to trigger a fever in the absence of any other infection. We can also consider a curve of threshold pairs for D and U above which a combined infection can trigger a fever.

Because Y indicates the presence of a fever resulting from any infection, we can treat Y as a function of both D and U. Y can be thought of also as a function of Z and G through their effect on D and U. Thus, we can write the potential outcome of Y for treatment z and sickle cell status g as $Y(D(z,g), U(z,g))$ and the observed outcome as $Y = Y(D(Z,G), U(Z,G))$. $Y(D(z,g), U(z,g))$ factors additively into two natural terms,

$$\underbrace{Y(D(z,g), U(z,g))}_{Y(z,g)} = \underbrace{\{Y(D(z,g), U(z,g)) - Y(0, U(z,g))\}}_{Y^m(z,g)} + \underbrace{Y(0, U(z,g))}_{Y^{nm}(z,g)}. \quad (1)$$

Y can be expressed as a potential outcome in terms of d and u or z and g. However, when the pyrogenic threshold depend on either z or g, Y can no longer be expressed as a potential outcome in terms of d and u alone. The protective mechanism of HbAS is understood to be complex, with some evidence supporting that part the protection conferred is by making the host more tolerant to infection, e.g., by increasing the pyrogenic threshold (see Section 3.3.2 for further discussion). The results in the remainder of the paper do not require that the pyrogenic be constant across levels of levels of z and g. When it is unambiguous, we use

the abbreviated notation below the "curly" braces in (1) to emphasize its dependence on z and g.

The statistical implication of a pyrogenic threshold is that $Y(d, u)$ is monotonic in d for all u. That is, $Y(d, u) \leq Y(d', u)$ for $0 \leq d \leq d'$. This ensures that the first term on the right hand side of (1) is non-negative. This term, $Y^m(z, g)$, can be interpreted as *malaria-attributable fever*, or a fever that would not have occurred had a malaria infection been absent. These include cases where D was high enough to trigger a fever in the absence of any other infection and also cases where D was high enough to trigger a fever in conjunction with a non-malaria infection. The second term on the right hand side, $Y^{nm}(z, g)$, is a fever that cannot be attributed to malaria. These are fevers that would have still occurred if malaria parasites were not present. However, when $D > 0$, $Y^m(Z, G)$ is generally not observable because it involves the counterfactual term $Y(0, U(Z, G))$. In the following section we will demonstrate that under a few realistic assumptions we can identify vaccine efficacy without directly observing the fevers (or deaths) attributable to malaria.

3.2.3 A Couple Final Pieces of Notation

With our potential outcome framework in place, we denote the complete data vector by

$$C = (X, Z, G, D, U, \{Y^k(z, g) : k = m, nm; z = 0, 1; g = 0, 1\}).$$

In reality, we only get to observe one potential outcome $Y(Z, G)$ even though all four exist and, as we have argued, Y^{nm} and Y^m cannot be distinguished with certainty so we only observe $O \subset C$. Finally, we may sometimes treat Y, D and U as $p \times 1$ vectors that correspond to fever status, parasite density, and non-malaria infectious load at p follow-up visits during a RCT.

3.3 Vaccine Efficacy: Definitions, Assumptions and Identification

3.3.1 Defining Vaccine Efficacy

We begin this section by defining a general population-level estimand for vaccine efficacy and then outline the assumptions required for identification when we cannot distinguish Y^m from Y^{nm}. We suppose that for each j, C_{ij} (and O_{ij}) are i.i.d. draws from an from an

unknown, site-specific target population distribution $P_j \in \mathcal{P}_j$ for all i. We suppose that the overall target population is an equally weighted mixture of the P_j and denote it $P \in \mathcal{P}$. We let $\mu_{zg}(P)$, $\mathcal{P} \to \mathbb{R}$ be a functional of P that depends on treatment and sickle cell status and takes the form $\mu_{zg}(P) = \mathbb{E}_P[f\{Y(z,g)\}]$ for P-measurable functions f that are (1) increasing and (2) linear. We use a superscript m to indicate that the functional is specific to malaria-attributable outcomes and nm to indicate that it is specific to non-malaria outcomes, for instance, $\mu_{zg}^m(P) = \mathbb{E}_P[Y^m(z,g)]$ when $f(y) = y$. Below we give a general definition of efficacy.

Definition 1 (Vaccine/Protective Efficacy). For a specified function f,

(i) Vaccine Efficacy is defined as $\tau(g) = 1 - \mu_{1g}^m(P)/\mu_{0g}^m(P)$ for $g = 0, 1$ and

(ii) Protective Efficacy of G is defined as $\nu(z) = 1 - \mu_{z1}^m(P)/\mu_{z0}^m(P)$ for $z = 0, 1$.

We are primarily interested in two simple functions f: $f(y) = y$ and when Y is a $p \times 1$ vector of fevers, $f(y) = \mathbf{1}^T y$. For $f(y) = y$, efficacy as defined in Definition 1 is the proportion reduction in risk of malaria-attributable fever or death were an individual to receive vaccination versus placebo. For $f(y) = \mathbf{1}^T y$, efficacy is defined as the proportion reduction in incidence of malaria-attributable fever were an individual to receive vaccination versus placebo. Since you can only die once, $f(y) = \mathbf{1}^T y$ is not applicable when Y indicates malaria-attributable death. By this same logic, $f(y) = y$ can be used for assessing the risk of malaria-attributable deaths over arbitrarily long follow-up. However, if we are interested in the risk of developing a malaria-attributable fever over a longer follow-up, during which a subject can have multiple fevers, the function we'd be interested in is $f(y) = \max y$ where $\max y$ is the maximum over the elements of a $p \times 1$ vector y. This function is not a linear function and thus vaccine efficacy against the risk of having at least one malaria-attributable fever over a long follow cannot be handled in this framework. The importance of the linearity of f will become evident shortly.

3.3.2 Identifying Vaccine Efficacy

In §2 we informally discussed the assumptions required to identify vaccine efficacy using MFD. Assumptions 1 - 5 formalize the assumptions that are summarized in the second column of Table 1 and provide the basis for the identification of τ proved in Proposition 1.

Assumption 1 (Additivity of μ). For all $z = 0, 1$ and $g = 0, 1$, $\mu_{zg}(P)$ can be decomposed linearly as

$$\mu_{zg}(P) = \mu_{zg}^m(P) + \mu_{gz}^{nm}(P).$$

Assumption 1 is guaranteed by (1) and the linearity of f. We mentioned above that evaluating the vaccine efficacy on the risk scale for malaria-attributable fevers over a long follow-up is problematic when using MFD to estimate τ. The risks of developing a malaria-attributable fever and a non-malaria fever do not satisfy Assumption 1 over long follow-ups because of the non-linearity of $f(y) = \max y$.

Assumption 2 (No Interaction / Independent Protective Pathways). $\tau(g)$ is constant over $g = 0, 1$ and $\nu(z)$ is constant over $z = 0, 1$. That is, $\tau := \tau(0) = \tau(1)$ and $\nu := \nu(0) = \nu(1)$.

The biological reasoning behind Assumption 2 is as follows. If a vaccine and genetic trait protect against malaria through independent pathways at different times in the parasite's life cycle, it is plausible that the vaccine will prevent the same fraction of fevers among those who have the genetic trait ($G = 1$) and those who do not ($G = 0$). Similarly plausible is that possession of the genetic trait will prevent the same fraction of fevers among those to whom the vaccine was administered ($Z = 1$) and to those it was not ($Z = 0$). The goal of pre-erythrocytic vaccines like RTS,S is to provoke an immune response that prevents the parasites from entering the liver, stopping the parasites from ever re-entering the bloodstream and causing clinical symptoms (Regules et al., 2011). In contrast, the sickle cell trait appears to protect against clinical symptoms by inhibiting the growth of parasites once they have re-entered the bloodstream from the liver and by making the host more tolerant to parasite infection (Ferreira et al., 2011; Taylor et al., 2012; Williams, 2011). This suggests that Assumption 2 is satisfied for pre-erythrocytic vaccines and the sickle cell trait.

Assumption 3 (No Interference). A subject's potential outcomes are functions of its vaccine and sickle cell status alone. That is $Y_{ij}(\mathbf{z}, \mathbf{g}) = Y_{ij}(z_{ij}, g_{ij})$.

Assumption 3 implies that the treatment and sickle cell status of an individual in the trial effects only their own outcome. This assumption can be rephrased in the context of infectious disease as stating that protection conferred by genetics or vaccine do not materially disrupt disease transmission. This assumption is plausible when considering individuals at two different sites $j \neq j'$ in a multi-site trial but is more complicated for individuals at the

same site who may live in close proximity. Stochastic simulation models of malaria vaccination using pre-erythrocytic and blood stage-vaccines have, however, found that transmission effects of such vaccines delivered as they would be in pediatric malaria vaccination programs were minimal (Penny et al., 2008, 2015). Less is known about the suitability of Assumption 3 with respect to the sickle cell variant.

Assumption 4 (Randomization / "as-if" Randomized). At each site $j = 1, \ldots, J$, vaccines are administered randomly and sickle cell status is distributed "as-if" random; that is,

$$(Z, G) \perp\!\!\!\perp (Y(z, g), Y^m(z, g), Y^{nm}(z, g), X)$$

for all $z, g \in \{0, 1\}^2$. Additionally, we assume that $Z \perp\!\!\!\perp G$.

The randomization of Z is ensured by the RCT. We assume that within each center, the sickle cell trait is distributed "as-if" random. Population stratification and linkage disequilibrium are common biological violations of the "as-if" random assumption of Mendelian genes (Kang et al., 2013). Briefly, population stratification is when a subgroups that differ on prognostic factors for developing malaria and other childhood illnesses also systematically differ in the prevalence of HbAS. If either the probability of inheriting HbAS or the distribution of prognostic factors is relatively homogenous within a study site, then population stratification would likely not be a material threat to the "as-if" random assumption about G. Linkage disequilibrium is the dependence of gene frequencies at two or more loci (Morton, 2001). If HbAS is in linkage disequilibrium with a gene that affects the risk of childhood illness this may threaten the validity of Assumption 4. Linkage disequilibrium with a gene that affects the risk of malaria is less problematic as this would not violate the exclusion restriction formalized in the next assumption. However, this would require a more delicate treatment of potential outcomes, e.g., we would need to consider the genes in linkage disequilibrium as having potential outcomes depending on G (see VanderWeele and Hernan (2013) for a discussion a related topic of "versions" of treatment).

Assumption 5 (Valid Mendelian Factor). G is a "valid Mendelian factor" in that it satisfies the following conditions:

(i) $\nu \neq 0$; and

(ii) $1 - \mu_{z1}^{nm}(P)/\mu_{z0}^{nm}(P) = 0$ for $z = 0, 1$.

Part (i) of Assumption 5 says that the Mendelian factor is *relevant* to malaria-attributable outcomes. A large body of literature on the protective properties of the sickle cell trait against malaria supports this assumption for HbAS. Several cohort studies have found evidence that HbAS has 30-50% efficacy against uncomplicated clinical malaria and multiple other case-control and cohort studies of Africa have estimated even greater efficacies against sever malaria cases of 70-90% (see Gong et al. (2013) for a list of several studies reporting the protective efficacy of HbAS). Assumption 5(ii) can be expressed in terms of potential outcomes by the restriction that U depends only on treatment status, $U(z, g) = U(z)$. The plausibility of Assumption 5(ii) is supported by evidence that the protection conferred by HbAS is "remarkably specific" to malaria, providing little protection to other childhood diseases (Williams et al., 2005).

The following proposition provides a simple, non-parametric identification strategy for treatment efficacy as defined in Definition 1 under Assumptions 1 - 5. We also make the standard assumptions of *consistency*, that $Y = Y(z, g)$ when $Z = z$ and $G = g$, and *positivity*, that the probability of treatment Z and the prevalence of G in each site are both bounded away from 0 and 1.

Proposition 1 (Nonparametric Identification). Suppose that Assumptions 1 - 5 are satisfied. Then the vaccine efficacy τ is identified from the observed data $\mathbf{O}$ as

$$\tau = 1 - \frac{J^{-1} \sum_{j=1}^{J} \left\{ \mathbb{E}_{X,j} \left[\mathbb{E}_{P_j}[f(Y) | X, Z = 1, G = 1] \right] - \mathbb{E}_{X,j} \left[\mathbb{E}_{P_j}[f(Y) | X, Z = 1, G = 0] \right] \right\}}{J^{-1} \sum_{j=1}^{J} \left\{ \mathbb{E}_{X,j} \left[\mathbb{E}_{P_j}[f(Y) | X, Z = 0, G = 1] \right] - \mathbb{E}_{X,j} \left[\mathbb{E}_{P_j}[f(Y) | X, Z = 0, G = 0] \right] \right\}}, \tag{2}$$

where $\mathbb{E}_{X,j}$ is the expectation over the marginal distribution of X implied by P_j.

Proof. The proof of Proposition 1 can be found in Appendix B.1 of the Supplemental Web Materials. □

3.3.3 Remarks

Proposition 1 still holds under a weaker version of Assumption 4 requiring only that Z and G are independent of potential outcomes (and of each other) *conditional* on baseline covariates X. Depending on how rich the set of baseline covariates X is, this weaker assumption may be more tenable in the presence of population stratification and linkage disequilibrium.

Equation (2) suggests a ratio estimator for τ that is remarkably similar to the Wald estimator used in Mendelian randomization studies (Wald, 1940; Burgess et al., 2017). Without covariates, the Wald estimator of the effect of a non-randomized exposure E on outcome Y using Mendelian gene G can be written as

$$\frac{\bar{Y}_{G=1} - \bar{Y}_{G=0}}{\bar{E}_{G=1} - \bar{E}_{G=0}} , \tag{3}$$

where $\bar{V}_{G=g}$ is the sample average of V for individuals with $G = g$. Both (2) and (3) involve ratios of averages differenced over G. The analogy between (2) and the Wald estimator in (3) is not by coincidence, but instead arises from a symmetry between the structures of Mendelian randomization with additive effects and Mendelian factorial design with multiplicative effects. More precisely, in Mendelian randomization, the association between the Mendelian gene and the outcome can be decomposed *multiplicatively* into the *additive* association of the gene, or instrument, on the exposure and the *additive* effect of the exposure on the outcome. In a Mendelian factorial design, the outcomes can be *additively* decomposed into disease-attributable outcomes and non-disease outcomes upon which the treatment and Mendelian factor have *multiplicative* effects. This simple structure of Mendelian factorial design can be seen clearly in the 2×2 table in Figure 2 of expected outcomes μ_{zg} for different combinations of z and g. Differencing over the columns and taking the ratio over the rows immediately yields $1 - \tau$. The η term in the top row can be thought of as the *spillover efficacy* of the vaccine against non-malaria outcomes. The term drops out when differencing over the columns.

An anonymous reviewer pointed out that segregation distortion and dynastic effects could possibly violate the identifying assumptions. For instance, children without the sickle cell trait are more likely to get symptomatic malaria and it has been demonstrated that malaria causes stunting in children (Kang et al., 2013), which might influence future response to vaccine. This would violate assumption 2. However, it has been observed that stunting and wasting is not associated with subsequent malaria infections (Gari et al., 2018) providing at least some evidence that this particular malaria-caused outcome does not modify vaccine efficacy by making individuals more susceptible to malaria. There is also partial evidence that dynastic effects, especially derived from non-transmitted maternal alleles may not materially threaten Assumption 4; Patel et al. (2016) show that maternal HbAS does not confer

protection against placental malaria, suggesting that it is not an unreasonable assumption that HbAS protection against malaria is passed from parent to child only through allele transmission and not through any other pathway, i.e., $G \leftarrow U \rightarrow Y$ in the right panel of Figure 1. As is the case with untestable assumptions, auxiliary evidence and partial empirical support is usually the best we can do to provide reassurance of their validity.

	g = 1	**g = 0**
z = 1	$\mu^{nm}(1-\eta) + \mu^m(1-\tau)(1-\nu)$	$\mu^{nm}(1-\eta) + \mu^m(1-\tau)$
z = 0	$\mu^{nm} + \mu^m(1-\nu)$	$\mu^{nm} + \mu^m$

Figure 2: 2×2 table for Mendelian factorial design. Each cell represents $\mu_{zg}(P)$ for all combinations of $(z, g) \in \{0, 1\}^2$, which is identified by the observed data $\mathbf{O}$ (Proposition 1). Differencing over the columns then taking the ratio over the rows yields $1 - \tau$. For notational clarity, we drop the subscript from μ_{00}. η is the "spillover efficacy" the vaccine may provide against non-malaria outcomes.

3.4   Robust Covariance Adjusted Estimation and Inference

We now propose a simple substitution estimator that allows for covariance adjustment and is robust to arbitrary misspecification of a model for $\mathbb{E}_{P_j}[f(Y) \,|\, Z, G, X]$. The estimation procedure closely resembles that which is developed in Rosenblum and van der Laan (2010) with a couple minor modifications to deal with the factorial structure of our identification procedure and the possibility that HbAS prevalence varies across sites in a multi-site RCT. The simple procedure is detailed below in Algorithm 1 and requires estimating a simple generalized linear model (GLM) with the R function `glm` in the `stats` package at most two times. In what follows, we suppose $f(y) = \mathbf{1}^T y$ for expository purposes. That is, we will focus on vaccine efficacy defined as the proportion reduction in the expected number of malaria-attributable fevers.

Algorithm 1 (Estimation of τ). The following substitution estimator is based on Rosenblum and van der Laan (2010). For clarity, we let $f(y) = \mathbf{1}^T y$.

---

1. Estimate $\mathbb{E}_{P_j}[f(Y) \,|\, Z, G, X]$ with `glm` using a canonical link function (depends on f).

   - Let the linear part include an intercept, main terms for Z and G, an, interaction $Z \times G$, and a site indicator (with possible interactions), for example,

     `mu_hat_0 <- predict(glm(f(Y) ~ S + X*G*Z - 1, family = poisson()))`

     where `S` is an indicator for belonging to site j.

   - Denote the resulting estimator when `S` $= j$ as $\hat{\mu}_{j,0}(Z, G, X)$.

2. If there is more than one site and the prevalence of G and sample sizes vary across sites, let `log_mu_hat` $= \log \hat{\mu}_{j,0}(Z_{ij}, G_{ij}, X_{ij})$, `p_G` $= (1/I_j)\sum_{i=1}^{I_j} G_{ij}$, and `w` $= n/(J \cdot I_j)$; update $\hat{\mu}_0$ as follows

   ```
   mu_hat_1 <- predict(glm(f(Y) ~ offset(log_mu_hat) + I(w*Z*G/p_G) +
                          I(w*Z*(1-G)/(1-p_G)) + I(w*(1-Z)G/p_G) +
                          I(w*(1-Z)*(1-G)/(1-p_G)) - 1,
                          family = poisson()))
   ```

   and denote the resulting estimator as $\hat{\mu}_{j,1}(Z, G, X)$.

3. Let $\mu_{zg}(P_n) = \frac{1}{J}\sum_{j=1}^{J} \frac{1}{I_j} \sum_{i=1}^{I_j} \hat{\mu}_{j,1}(z, g, X_{ij})$.

4. Construct the plug-in MFD estimator

$$\hat{\tau} = 1 - \frac{\mu_{11}(P_n) - \mu_{10}(P_n)}{\mu_{01}(P_n) - \mu_{00}(P_n)}.$$

The estimator $\hat{\tau}$ returned by Algorithm 1 is a special case of a target maximum likelihood estimator (TMLE) (Van der Laan and Rose, 2011). A straightforward modification of Theorem 1 in Rosenblum and van der Laan (2010) yields that $\hat{\tau}$ is consistent and asymptotically normal under mild regularity conditions even when the working model for the conditional

expectation is misspecified. Other initial working models $\hat{\mu}_0$ may be used as long as they satisfy certain restrictions on how data-adaptive they are (Van der Laan and Rose, 2011). The weight terms w are important because we assume that the target population P is an equally weighted mixture of the site-specific populations P_j but allow for study designs with different sample sizes across sites. The weights ensure that $\hat{\mu}_1$ solves the efficient influence function estimating equation (see Appendix B.4 of the Supplemental Web Materials). Before we state the result, we introduce some some important notation and definitions.

Working Models and their Limits:

Let the maximum likelihood parameter estimates from steps 1 and 2 of Algorithm 1 be $\boldsymbol{\beta}_n^{(0)}$ and $\boldsymbol{\beta}_n^{(1)}$, respectively, and define $\boldsymbol{\beta}_n = [\boldsymbol{\beta}_n^{(0)}, \boldsymbol{\beta}_n^{(1)}]$. Now, recall that P_j is the true, unknown data generating distribution for site j. We can decompose its corresponding density p_j as follows

$$p_j = p_j(X)p(Z)p_j(G)p_j(Y \,|\, Z,\, G,\, X) \tag{4}$$

where $p_j(G)$ is the probability a child has the sickle cell trait in site j. $p(Z)$ is assumed to be known, e.g., $p(Z) = 1/2$ in a balanced trial. The density for the overall target population distribution can be written as $p = p(S = j)p_j$. $P_{j,n}$ is our estimate of P_j where $p_{j,n}$, the density of $P_{j,n}$, can be decomposed similarly as

$$p_{j,n} = p_{j,n}(X)p(Z)p_{j,n}(G)p_{j,\boldsymbol{\beta}_n}(Y \,|\, Z,\, G,\, X) \tag{5}$$

where $p_{j,n}(X)$ is the empirical distribution of X at site j, $p_{j,n}(G)$ is the observed prevalence of the sickle cell trait among children in enrolled in the study at site j, and $p_{j,\boldsymbol{\beta}_n}(Y \,|\, Z,\, G,\, X)$ is the estimated parametric working model for the conditional distribution of Y from steps 1 and 2 in Algorithm 1. The resulting estimate of the overall target population distribution can be written as $p_n = p(S = j)p_{j,n}$. We assume that the number of site are fixed, that they are representative of the population of interest, and that the sites carry equal weight in the population they represent – $p(S = j) = 1/J$ for all j – but that the sample sizes I_j might be different due to study design rather than just chance. Hence, the observations that make up the empirical distribution are weighted by $n/(J \cdot I_j)$. We assume that the site-level sample sizes I_j grow at the same rate as $n \to \infty$ and so it follows that $p_{j,n}(X) \overset{a.s.}{\to} p_j(X)$ as $n \to \infty$ by the Gilvenko-Cantelli Theorem and an application of the strong law of large numbers

18

gives us that $p_{j,n}(G) \overset{a.s.}{\to} p_j(G)$ for all $j = 1, \ldots, J$ as $n \to \infty$. Let $P_{j,\infty} = \lim_n P_{j,n}$ and write its density $p_{j,\infty}$ as

$$p_{j,\infty} = p_j(X)p(Z)p_j(G)p_{j,\boldsymbol{\beta}}(Y \mid Z, G, X) \tag{6}$$

where $\boldsymbol{\beta} = [\boldsymbol{\beta}^{(0)}, \boldsymbol{\beta}^{(1)}]$ are the maximizers of the expected log-likelihoods of the GLMs in steps 1 and 2 where the expectation is taken over P, if such maximizers exist (see Rosenblum and van der Laan (2010) for further discussion of the existence of $\boldsymbol{\beta}$). We write $p_\infty = p(S = j)p_{j,\infty}$. When $\boldsymbol{\beta}$ exists, the conditions given in Proposition 2 are sufficient for $\boldsymbol{\beta}_n$ to converge to $\boldsymbol{\beta}$ in probability (Rosenblum and van der Laan, 2009). Note that unless the working parametric model for the conditional mean is correctly specified, P_∞ will not equal P in general.

For notational convenience, we let $\mathbb{P}_j$ and $\mathbb{P}_{j,n}$ be the expectation operators over P_j and $P_{j,n}$, respectively, and define $\mathbb{P}$ and $\mathbb{P}_n$ similarly. For example, we can write $\mu_{zg}(P_n)$ from step 3 of Algorithm 1 as $\mathbb{P}_n \hat{\mu}_{j,1}(z, g, X)$. Alternatively, we can write $\mu_{zg}(P_{j,n}) = \mathbb{P}_{j,n} \hat{\mu}_{j,1}(z, g, X)$ and $\mu_{zg}(P_n) = (1/J) \sum_j \mu_{zg}(P_{j,n})$.

Efficient Influence Functions:

For an arbitrary distribution $Q \in \mathcal{P}$, the *efficient influence function* (EIF) for $\mu_{zg}(Q)$ when the Q_j and q_j are known can be written as $\varphi_{zg}(Q) = J^{-1} \sum_{j=1}^{J} \varphi_{zg}(Q_j)$ where

$$\varphi_{zg}(Q_j)(O) = \frac{\mathbb{1}(Z = z)\mathbb{1}(G = g)(f(Y) - \mathbb{E}_{Q_j}[f(Y) \mid Z = z, G = g, X])}{q_j(G = g)q(Z = z)}$$

$$+ \mathbb{E}_{Q_j}[f(Y)], Z = z, G = g, X, S] - \mu_{zg}(Q_j), \tag{7}$$

for $z, g \in \{0, 1\}^2$. We will define $\mu_z(Q) := \mu_{z1}(Q) - \mu_{z0}(Q)$ and $\varphi_z(Q) := \varphi_{z1}(Q) - \varphi_{z0}(Q)$ for $z = 0, 1$. Standard calculations verify that $\varphi_z(Q)$ is the efficient influence function for $\mu_z(Q)$ by demonstrating that $\varphi_z(Q)$ can be expressed as a pathwise derivative of $\mu_z(Q_\epsilon)$ where Q_ϵ is a parametric submodel of Q such that $Q = Q_{\epsilon=0}$ (Kennedy, 2016). $\varphi_z(Q)$ is said to be a pathwise derivative of $\mu_z(Q_\epsilon)$ if $\mathbb{E}_Q[\varphi_z(Q)S_\epsilon] = \partial \mu_z(Q_\epsilon)/\partial \epsilon|_{\epsilon=0}$ where S_ϵ is the score function of the parametric submodel. Finally, we project $\varphi_{zg}(Q)$ onto the tangent space of the model for G because we have to (consistently) estimate the prevalence of G at each site j (Van der Laan and Rose, 2011, p. 96). We let $\varphi_{zg}^*(Q) = \varphi_{zg}(Q) - \mathbb{E}_P[\varphi_{zg}(Q) \mid G]$ and define $\varphi_z^*(Q) := \varphi_{z1}^*(Q) - \varphi_{z0}^*(Q)$. We can now state the consistency and asymptotic normality result for $\hat{\tau}$.

Proposition 2 (Consistency and Asymptotic Normality). In addition to Assumptions 1 - 5, suppose that the number of sites J is fixed, the I_j grow at the same rate as n, and the maximizers $\boldsymbol{\beta} = [\boldsymbol{\beta}^{(0)}, \boldsymbol{\beta}^{(1)}]$ exist. Also, suppose that $\|\boldsymbol{\beta}\|_\infty < M$ for pre-specified $M < \infty$. Finally, let (X, Y) be bounded and the terms in the linear parts of the GLMs in steps 1 and 2 of Algoritm 1 be bounded functions on compact subsets of $\{0, 1\}^2 \times \mathbb{R}^d$. Then, $\hat{\tau}$ is consistent and $\sqrt{n}(\hat{\tau} - \tau)$ convergence in distribution to a Gaussian with mean 0 and variance

$$\sigma^2 = \tilde{C} \cdot \mathbb{P} \left[\frac{\mu_1(P)}{\mu_0(P)^2}(\varphi_0^*(P_\infty)(O) - \mathbb{P}\varphi_0^*(P_\infty)(O)\} - \frac{1}{\mu_0(P)}\{\varphi_1^*(P_\infty)(O) - \mathbb{P}\varphi_1^*(P_\infty)(O)\} \right]^2 ,$$

where $\tilde{C} = (nJ^{-2})\sum_{j=1}^J I_j^{-1} \geq 1$ is correction factor for when study samples differ in size across sites due to design rather than chance. When the prevalence of G does not vary between sites and the study is balanced, i.e., $I_j = n/J$ for all j, then step 2 of Algorithm 1 can be skipped. Step 2 may also be skipped if it is a single-site trial. Furthermore, if the working model for $\mathbb{E}_P[f(Y) \mid Z, G, X, S]$ is correctly specified then σ^2 achieves the semiparametric efficiency bound.

Proof. A sketch of the proof of Proposition 2 can be found in Appendix B.4 of the Supplemental Web Materials. □

Remark 1. We remarked in §3.3.2 that Proposition 1 still holds under a weaker version of Assumption 4 where the independence is conditional on baseline covariates X. Proposition 2 also holds under these weaker assumptions as long as either the working model for the conditional expectation or the conditional assignment models for G and Z are correctly specified.

Variance Estimator:

Notice that φ_z^* is defined as a projection of φ_z into a lower dimensional subspace and will thus have smaller variance. With that in mind, we can use $\varphi_z(P_n)$ to construct a conservative plug-in estimator of σ^2. In Appendix B.4 of the Supplemental Web Materials we show that P_n solves the EIF estimating equations, that is, $\mathbb{P}_n\varphi_z(P_n)(O) = 0$ for $z = 0, 1$. It follows then that the plug-in estimator of the scale variance σ^2 can be written simply as

$$n \cdot \widehat{\text{var}}(\hat{\tau}) = \tilde{C} \cdot \mathbb{P}_n \left[\varphi_0(P_n)(O)\frac{\mu_1(P_n)}{\mu_1(P_n)^2} - \varphi_1(P_n)(O)\frac{1}{\mu_0(P_n)} \right]^2 \tag{8}$$

With this variance estimator we can now use Proposition 2 to conduct inference on and construct confidence intervals for $\hat{\tau}$.

Naive Estimator:

For simplicity, let's assume that we have balanced sites, i.e., $I_j = n/J$ for all j. Then, had we assumed that all fevers with any parasitemia were malaria-attributable fevers, we might considered the following *naive estimator*

$$\hat{\tau}_0 = 1 - \frac{\mathbb{P}_n \hat{\mu}_0(1, G, X, S)}{\mathbb{P}_n \hat{\mu}_0(0, G, X, S)}. \tag{9}$$

The naive estimator corresponds with standard estimates of VE with respect to a commonly used secondary case definition of the presence of a fever ($Y = 1$) and any positive parasite density ($D > 0$) (Olotu et al., 2013). We will see in the following section that the estimator is systematically biased but can be combined with our MFD estimator $\hat{\tau}$ to construct an estimator that outperforms both estimators on their own.

3.5 Simulation Study: Comparison to Naive Identification Strategy

In this section we investigate the performance of our proposed MFD estimator $\hat{\tau}$ and compare it to that of the naive estimator $\hat{\tau}_0$ that assume all fevers with any parasitemia are malaria fevers. As in the previous section, we consider $f(y) = \mathbf{1}^T y$ and thus τ is the proportion reduction in expected number of malaria-attributable fevers. We consider a single-site RCT $J = 1$ with equal sized vaccine and placebo arms and the prevalence of HbAS set to 20% based on existing estimates of the prevalence in sub-Saharan Africa (Ter Kuile et al., 2003; Elguero et al., 2015).

We simulate the number of malaria-attributable fevers and non-malaria fevers from negative binomial distributions in a single year of follow up for each individual. Evidence suggests that the negative binomial distribution fits the empirical distribution of the number of clinical malaria events an individual experiences well (Olotu et al., 2013). We can also see from (1) that Y^m and Y^{nm} are negatively dependent – a fever cannot be both attributable to malaria and have been present in the absence of malaria. To model this dependence structure, we use a Gaussian copula with negative dependence parameter $\rho = -0.1$ (Genest and Neslehová, 2007). The negative dependence is modest when fevers are rare but will be

more pronounced in areas where fever risk is higher, e.g., villages with poor sanitation. We suppose there is a single observed covariate X that enters the conditional mean function for both $\mathbf{1}^T Y^m$ and $\mathbf{1}^T Y^{nm}$ and that there is unobserved heterogeneity in the conditional means between individuals. The generating distributions were calibrated so that the average number of any-cause fevers is 1.5 per child-year. We also calibrated the generating distributions to achieve different levels of case specificity, which we define formally below.

Two different sample sizes were chosen to assess how performance improved asymptotically: 1000 subjects per trial arm, i.e., $n = 2000$, and 2500 subjects per trial arm, i.e., $n = 5000$. For $n = 2000$, we consider $5 \times 2 \times 2$ different combinations of vaccine efficacy $\tau \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$, protective efficacy of the Mendelian factor $\nu \in \{0.3, 0.5\}$, and specificities $s \in \{0.5, 0.8\}$. Formally, we define specificity s here as the expected number of malaria-attributable fevers divided by the expected number of fever with any parasitemia and of any cause under placebo,

$$s = \frac{\mathbb{E}_P\left[\mathbf{1}^T Y^m(0, G)\right]}{\mathbb{E}_P\left[\mathbf{1}^T Y(0, G)\right]} = \frac{p(G=1)\mu_{01}^m + p(G=0)\mu_{00}^m}{p(G=1)\mu_{01} + p(G=0)\mu_{00}}. \tag{10}$$

The specificity choices are motivated by Mabunda et al. (2009), which estimated the specificity of standard case definitions using > 0 and > 2500 parasites per μl cutoffs for children under five years of age to be roughly 50% and 80%, respectively. The strength of the Mendelian factor was calibrated to the aforementioned estimates of the protective efficacy of HbAS. We consider the same combinations for $n = 5000$ except only for the weaker protective efficacy setting. The spillover efficacy η was assumed to be zero. Finally, we allow for non-constant vaccine and Mendelian protective efficacy – both τ_i and ν_i are log-normally distributed with mean τ and ν, respectively, and standard deviation approximately 0.05 giving coefficients of variation of about 8-17%. Rather than conferring complete protection to a fraction of the vaccinated subjects, all vaccinated subjects receive partial protection. This is consistent with evidence that RTS,S is a "leaky" vaccine, providing at least partial protection to all recipients of the vaccine (Moorthy and Ballou, 2009). Each setting was simulated $N_{sim} = 5000$ times. More details of the simulation settings can be found in Appendix B.5 of the Supplemental Web Materials. R code to reproduce the simulation results can be found in the Supplementary Materials.

We use Poisson regression for the initial working model estimate in Algorithm 1 and

because $J = 1$, we skip step 2. The results of the simulation study are summarized in the following two tables. In Table 2 we compare the absolute proportional bias and root mean squared error (RMSE) of the $\hat{\tau}$ (MFD) and $\hat{\tau}_0$ (Naive). For $n = 2000$, the MFD estimator has very good bias properties, with proportional bias less than 5% for most settings. The RMSE increases for Mendelian genes that are less protective and as the specificity decreases. The only area of poor bias performance is when the Mendelian gene is weakly protective ($\nu = 0.3$) and the specificity is low (0.5). However, although the estimator in this setting has high variance the bias and power at $\tau = 0.7$ are reasonably adequate. Like the presence of small sample bias and high variance for weak instruments in instrumental variable analysis (Imbens and Rosenbaum, 2005), the poor performance appears to be a small sample property as the performance for the weak Mendelian gene, low specificity setting improves notably for $n = 5000$.

The bias in the naive estimator only varies over the different specificity settings and does not improve as the sample size grows. In fact, because there is no spillover efficacy, the proportional absolute bias is equal to $1 - s$. You can see from the table that the RMSE is driven almost entirely by the bias component for the naive estimator and it actually increases in absolute terms as the vaccine efficacy increases. These properties lead to very poor coverage of confidence intervals derived from the naive estimator. That one should expect efficacy estimates to be biased by as much as 20% when specificity is at the level recommended by the WHO should be cause for concern.

Table 3 below compares the coverage of two-sided 95% confidence intervals and the power against the two-sided alternative at 5% significance for the MFD and naive procedures. The MFD confidence interval has correct or conservative coverage and decent power in even some of the more unfavorable settings. Even in the small sample, weakly protective Mendelian gene, and low specificity setting the power is not negligible at higher vaccine efficacies. Because the naive estimator is systematically biased, the coverage properties are extremely poor in all settings. The variance of the naive estimator tends to be far smaller than the MFD estimator. This fact coupled with the systematic bias leads to high power and low coverage across all settings.

Even in settings where the mean proportional bias is high, the MFD estimator appears to have the desirable property of being median unbiased. Figure 3 demonstrates this for

Table 2: Proportional absolute bias and root mean squared error (RMSE) of MFD and naive estimators using $N_{sim} = 5000$ simulations.

| | | Specificity = 0.8 | | | | Specificity = 0.5 | | | |
| | | Prop. \|Bias\| | | RMSE | | Prop. \|Bias\| | | RMSE | |
ν	τ	MFD	Naive	MFD	Naive	MFD	Naive	MFD	Naive
$n = 1000(\times 2)$									
0.50	0.30	0.03	0.20	0.15	0.07	0.12	0.50	0.29	0.15
	0.40	0.02	0.20	0.14	0.08	0.07	0.50	0.27	0.20
	0.50	0.02	0.20	0.12	0.10	0.05	0.50	0.25	0.25
	0.60	0.02	0.20	0.11	0.12	0.04	0.50	0.24	0.30
	0.70	0.01	0.20	0.10	0.14	0.02	0.50	0.22	0.35
0.30	0.30	0.12	0.20	0.36	0.07	0.59	0.50	5.47	0.15
	0.40	0.11	0.20	0.30	0.09	0.32	0.50	4.54	0.20
	0.50	0.08	0.20	0.29	0.10	0.21	0.50	5.80	0.25
	0.60	0.06	0.20	0.24	0.12	0.18	0.50	4.61	0.30
	0.70	0.03	0.20	0.21	0.14	0.11	0.50	1.69	0.35
$n = 2500(\times 2)$									
0.30	0.30	0.06	0.20	0.18	0.06	0.18	0.50	0.37	0.15
	0.40	0.04	0.20	0.17	0.08	0.13	0.50	0.34	0.20
	0.50	0.02	0.20	0.14	0.10	0.08	0.50	0.30	0.25
	0.60	0.02	0.20	0.13	0.12	0.06	0.50	0.29	0.30
	0.70	0.01	0.20	0.12	0.14	0.04	0.50	0.27	0.35

Table 3: Coverage of two-sided 95% confidence interval and power against two-sided alternative at 5% significance level of MFD and naive estimators using $N_{sim} = 5000$ simulations.

| | | Specificity = 0.8 | | | | Specificity = 0.5 | | | |
| | | Coverage | | Power | | Coverage | | Power | |
ν	τ	MFD	Naive	MFD	Naive	MFD	Naive	MFD	Naive
$n = 1000(\times 2)$									
0.50	0.30	0.95	0.50	0.56	1.00	0.96	0.00	0.29	0.99
	0.40	0.95	0.17	0.79	1.00	0.96	0.00	0.44	1.00
	0.50	0.96	0.01	0.94	1.00	0.96	0.00	0.59	1.00
	0.60	0.96	0.00	0.99	1.00	0.96	0.00	0.73	1.00
	0.70	0.96	0.00	1.00	1.00	0.97	0.00	0.84	1.00
0.30	0.30	0.95	0.50	0.30	1.00	0.95	0.00	0.19	0.99
	0.40	0.95	0.16	0.44	1.00	0.96	0.00	0.25	1.00
	0.50	0.96	0.01	0.57	1.00	0.96	0.00	0.33	1.00
	0.60	0.96	0.00	0.74	1.00	0.97	0.00	0.42	1.00
	0.70	0.96	0.00	0.86	1.00	0.98	0.00	0.51	1.00
$n = 2500(\times 2)$									
0.30	0.30	0.95	0.11	0.45	1.00	0.95	0.00	0.26	1.00
	0.40	0.96	0.00	0.68	1.00	0.96	0.00	0.38	1.00
	0.50	0.95	0.00	0.87	1.00	0.96	0.00	0.52	1.00
	0.60	0.96	0.00	0.96	1.00	0.97	0.00	0.66	1.00
	0.70	0.96	0.00	0.99	1.00	0.97	0.00	0.77	1.00

$\tau = 0.3, 0.5$, and 0.7 in all six settings detailed in Tables 2 and 3. The dashed lines indicates the true value of τ, the boxes represents the IQRs, and the whiskers indicate the 5% and 95% quantiles of the simulated estimates. The diamonds indicate the means and the vertical solid lines the medians. The worsening performance of the MFD as the Mendelian gene weakens is clear (settings 2 vs. 1, 4 vs. 3, and 6 vs. 5) as is the improvement in the weak Mendelian gene settings as the sample size grows (settings 5 vs. 3 and 6 vs. 4). The naive estimates have much lower variance but are systematically mean and median biased downward. The MFD estimates are also mean biased downward. Although the estimator is asymptotically normal, the ratio of means that are jointly asymptotically normal may have a peculiar non-normal form in finite samples which may explain this particular pattern of bias (Marsaglia et al., 2006).
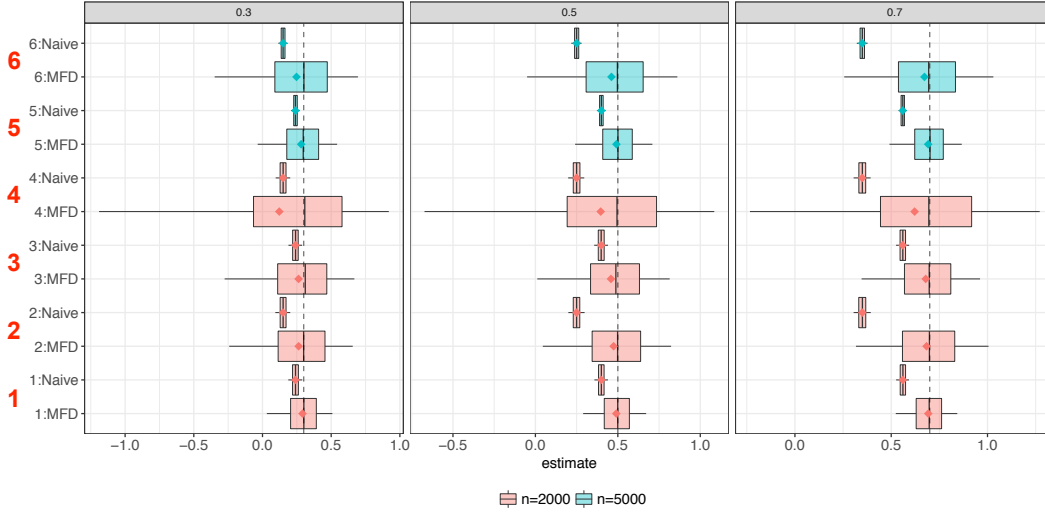

Figure 3: Distributions of simulated MFD estimator $\hat{\tau}$ and naive estimator $\hat{\tau}_0$ over several settings and $\tau = 0.3, 0.5, 0.7$. Setting **1**: strong factor ($\nu = 0.5$), high specificity ($s = 0.8$); setting **2**: strong factor, low specificity ($s = 0.5$); settings **3** and **5**: weak factor ($\nu = 0.3$), high specificity; settings **4** and **6**: weak factor, low specificity. Dashed lines indicate true efficacy, boxes indicate IQRs, whiskers are 5% and 95% quantiles, diamonds are mean estimate, and vertical solid lines are median estimates.

3.6 Improved Estimators: Leveraging the Identifying Assumptions

3.6.1 A Simple Correction to the Naive Estimator

The observation that the proportional absolute bias of $\hat{\tau}_0$ is equal to $1 - s$ in Table 2 for all settings comes from the fact that, when the sample sizes are balanced across sites, an immediate application of Theorem 1 of Rosenblum and van der Laan (2010) yields that $\hat{\tau}_0$ is a consistent estimator of $s\tau + (1-s)\eta$. When samples are small or the Mendelian gene is only weakly protective, we observed that the MFD estimator will be less powerful and may suffer from small sample bias. In such settings, the asymptotic limit of $\hat{\tau}_0$ suggests a simple correction to the naive estimator: dividing by s. Call this the *s-corrected estimator*, which is consistent for τ when $\eta = 0$ and s is known or consistently estimated itself. If instead we only have a $1 - \beta$ confidence interval for s, $\mathcal{C}_\beta$, then we can still construct a valid confidence interval for τ using the s-corrected estimator. Let $\text{CI}_{s,\alpha+\beta}$ be a $1 - \alpha - \beta$ confidence interval for τ constructed using the s-corrected estimator and define the s-corrected $1 - \alpha$ confidence interval $\text{CI}_{corr,\alpha} = \bigcup_{s \in \mathcal{C}_\beta} \text{CI}_{s,\alpha+\beta}$. Berger and Boos (1994) show that one can construct valid p-values by maximizing a non-pivotal p-value over the confidence set of a nuisance parameter. This result is easily inverted, providing a procedure to construct valid confidence intervals from which it follows that $\text{CI}_{corr,\alpha}$ will have the correct (conservative) coverage for τ. However, as we mentioned earlier in §3.2.1, estimating and conducting inference about s is challenging. However, we can still make use of the naive estimator even when we do not know s.

3.6.2 The Best of Both Worlds? Combining MFD and Naive Estimators

The naive estimator and the MFD estimator have complementary strengths and different weaknesses – $\hat{\tau}_0$ tends to be more efficient but is asymptotically biased and $\hat{\tau}$ is consistent but has higher variance and requires larger sample sizes. The naive estimator also has the nice feature of being a consistent lower bound of τ as long as $\eta \leq \tau$ – it is very plausible that a well designed vaccine will have a higher vaccine efficacy than spillover efficacy. Additionally, we have the logical constraint that $\tau \leq 1$ since a treatment efficacy greater than one would imply that it would be possible to have a negative number of malaria-attributable fevers, a scenario eliminated by the assumption that $Y(d, u)$ is monotonically increasing in d for

all levels u. The following proposition constructs an estimator that uses these upper and lower bounds to improve the performance of $\hat{\tau}$ in difficult settings, e.g., small samples, weak Mendelian factor, and low specificity.

Proposition 3 (Bounded Estimator). Suppose that $\eta \leq \tau$. Let $\hat{\tau}_0$ be the naive estimator and let

$$L_\alpha = \hat{\tau} - \Phi(1 - \alpha)\widehat{\mathrm{var}}(\hat{\tau})^{1/2}$$
$$L_{0,\alpha} = \hat{\tau}_0 - \Phi(1 - \alpha)\widehat{\mathrm{var}}(\hat{\tau}_0)^{1/2} .$$

Define the upper confidence bounds U_α and $U_{0,\alpha}$ similarly. Then (i)

$$\hat{\tau}_{bnd} = \min\left[1, \max\left\{\hat{\tau}, L_{0,\tilde{\alpha}}\right\}\right] \tag{11}$$

is consistent for τ when $\tilde{\alpha}$ is bounded away from 1; and (ii), for $0 \leq \alpha_0 \leq \alpha/2$

$$\mathrm{CI}_{bnd,\alpha} = \left[\max\{L_{\alpha/2-\alpha_0}, L_{0,\alpha_0}\}, \min\left\{1, U_{\alpha/2}\right\}\right] \tag{12}$$

is an asymptotically valid $1 - \alpha$ confidence interval for τ.

Proof. The proof of Proposition 3 can be found in Appendix B.3 of the Supplemental Web Materials. □

Because the naive estimator has much smaller variance than the MFD estimator, one will likely choose α_0 and $\tilde{\alpha}$ to be much smaller than α. In general, it makes sense to choose $\tilde{\alpha} = \alpha_0$ to ensure that $\hat{\tau}_{bnd} \in \mathrm{CI}_{bnd,\alpha}$.

The lower bound used to construct $\hat{\tau}_{bnd}$ acts like a high probability, stochastic lower bound to τ while 1 is an exact upper bound. When the Mendelian gene is weak, the denominator in $\hat{\tau}$ can sometimes be very close to zero, leading to unreliable estimates of vaccine efficacy. The bounded estimator is designed to mitigate the effect of these cases while keeping the median unbiasedness of $\hat{\tau}$ intact. Figure 4, illustrates how this plays out in a the setting where the $\hat{\tau}$ performs poorly with $n = 2000$, a weak Mendelian factor, low specificity, and $\tau = 0.5$. The figure shows the estimated densities and means of the naive estimator (blue, dash), MFD estimator (gray, dot), and bounded estimator (green, dot-dash) over 5000 simulations. The true vaccine efficacy is indicated by the black solid line, which lines up almost exactly with

the green dot-dashed line of the bounded estimator. As expected, $\hat{\tau}$ is much more variable and less biased than $\hat{\tau}_0$, but still materially biased in this setting. The bounded estimator $\hat{\tau}_{bnd}$ is nearly mean unbiased. You can see how it achieves this by "clumping" unreliable MFD estimates near the stochastic lower bound $\hat{\tau}_0$ and the exact upper bound 1.
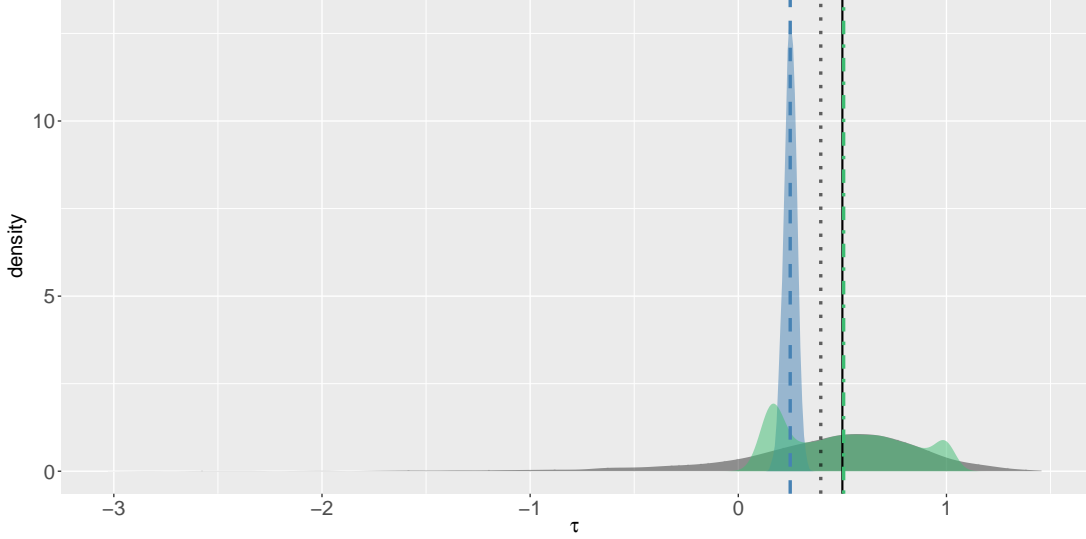

Figure 4: Densities and means of naive estimator (blue, dash) MFD estimator (gray, dot), and bounded estimator (green, dot-dash); True vaccine efficacy (black, solid). For the setting with $n = 1000(\times 2)$, $\tau = 0.5$, $\nu = 0.3$, spec. $= 0.5$, and $\tilde{\alpha} = 0.001$

We also assess how well $\hat{\tau}_{bnd}$ performs in the same setting as in Figure 4 but with an even smaller sample size ($n = 1000$). Using $\alpha = 0.05$, $\alpha_0 = 0.001$, and $\tilde{\alpha} = 0.001$, we find that $\hat{\tau}_{bnd}$ has substantially improved bias and RMSE while $\text{CI}_{bnd,\alpha}$ has very favorable power while maintaining the correct (conservative) coverage. In Figure 5, we compare the absolute proportional bias (left panel) and RMSE (right panel) of the three estimators. Absolute proportional bias and RMSE values larger than 1 are not shown. The bounded estimator uniformly outperforms both the MFD and naive estimators in terms of bias and has performance comparable to that of the MFD estimator estimated on a sample five times as large ($n = 5000$) for $\tau = 0.5, 0.6,$ and 0.7. The RMSE of $\hat{\tau}_{bnd}$ shows large improvements over the $\hat{\tau}$ and is comparable to the RMSE of $\hat{\tau}$ estimated in the same setting on a sample size of $n = 5000$ for all values of τ considered.

The power and coverage properties of $\text{CI}_{bnd,\alpha}$ give the clearest picture of how the comple-
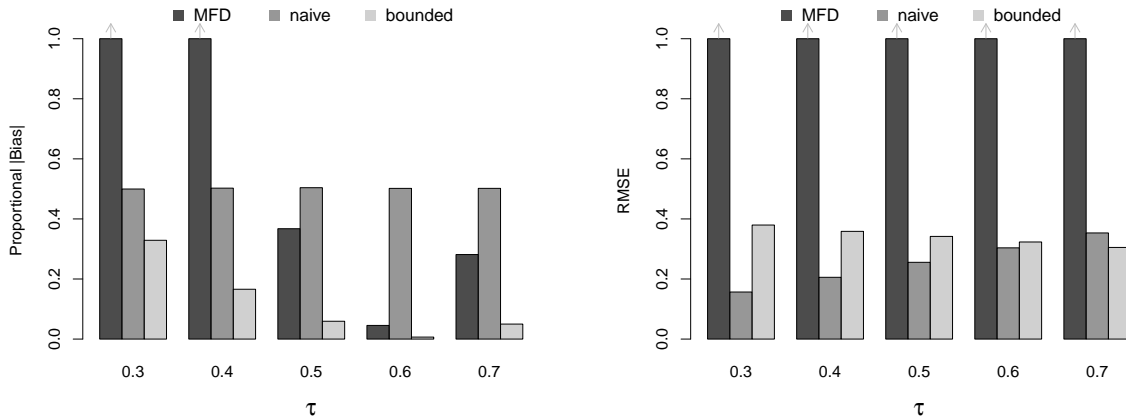
Figure 5: Absolute proportional bias (left panel) and RMSE (right panel) for MFD, naive, and bounded estimators with sample size $n = 1000$ and $N_{sim} = 5000$ simulations. For bias and RMSE values above 1, only a maximum of 1 is shown. The actual absolute proportional bias values for the MFD estimator are 2.21 and 1.56 for $\tau = 0.3$ and 0.4, respectively. The actual RMSE values for the MFD estimator from left to right are $18.58, 67.90, 32.29, 8.42,$ and 6.75.

mentary strengths of the MFD and naive procedures are retained by the bounded procedure. In the left panel of Figure 6, we see that $\text{CI}_{bnd,0.05}$ has only marginally more conservative coverage than the MFD confidence interval $[L_{0.025}, U_{0.025}]$ (right panel) while retaining the favorable power properties of the naive confidence interval $[L_{0,0.025}, U_{0,0.025}]$ (left panel). The dashed line in the right panel indicates 95% coverage.

Importantly, the superior performance of the bounded estimator in this particular setting does not appear to come at the expense of performance in the settings where $\hat{\tau}$ generally does well, improving on the $\hat{\tau}$ in almost all settings investigated in the simulation study.

4 Discussion

The strategy that we've developed for identification and estimation of malaria vaccine efficacy does not rely on the explicit definition of an inexact, but observable case definition. In short, we propose separating the approach to identifying VE into two distinct steps: first, define a gold-standard case that may be unobservable and then, identify VE using a strategy
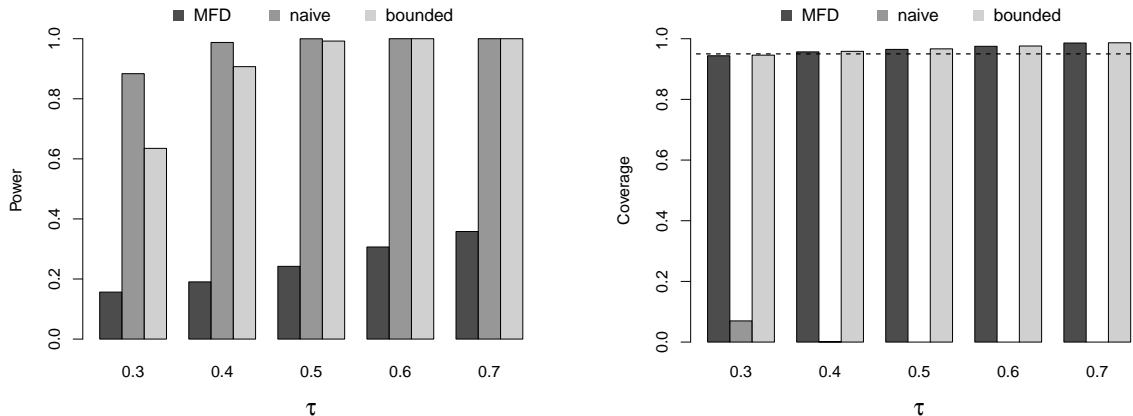
Figure 6: Power against two-sided alternative at 5% significance (left panel) and coverage of 95% confidence interval (right panel) for MFD, naive, and bounded estimators with sample size $n = 1000$ and $N_{sim} = 5000$ simulations. The dashed black line in the right panel indicates 95% coverage.

that doesn't require overt observation of these gold-standard cases. Our gold-standard case definition is one that is 100% specific and sensitive. It is precisely defined using a potential outcome framework. As we've noted, these cases are not distinguishable from observed data alone. Regardless, we demonstrate that with Mendelian factorial design we are able to leverage genetic variation in an analogous fashion to Mendelian randomization studies to identify VE with respect to this exact, but unobservable, case definition.

In observational studies, *evidence factors* are defined as several approximately independent tests of the same hypothesis, each of which depend on different assumptions about bias from non-random treatment assignment (Rosenbaum, 2011). In the presence of bias, these tests can be thought of as independent pieces of evidence in that the violation of assumptions underlying one test does not imply the other tests are similarly biased. Like an evidence factor, the MFD estimate can provide an additional piece of evidence in vaccine efficacy studies that relies on a different set of assumptions than the methods currently in use. For instance, the naive estimator assumes all fevers with any parasitemia are malaria-attributable, corresponding to a commonly used secondary case definition. Although not independent like true evidence factors, we demonstrated that the naive estimator and the MFD estimator can be combined to construct a *bounded* estimator that outperforms both estimators when

31

used on their own. In particular, the bounded estimator provides significant improvements when the Mendelian gene is only weakly protective, a challenging setting similar to the weak instrument setting in IV studies.

There is evidence that HbAS is to moderately protective against uncomplicated malaria and highly protective against severe malaria illness suggesting that MFD may be particularly useful for estimating VE against severe malaria, whose symptoms are not specific and overlap significantly with symptoms of other severe childhood comorbidities (Bejon et al., 2007).

The performance of the combined estimator across a range of simulation settings with sample sizes that are similar to those found in phase II and III clinical trials is promising. The results suggests that, if feasible, it would be prudent to begin collecting subject data on inherited hemoglobinopathies and other genetic traits that provide protection against clinical malaria, such as the sickle cell trait. Estimating the efficacy of prevention strategies that target transmission directly, such as a malaria transmission-blocking vaccine (Wu et al., 2015) and insecticide-treated bed nets (Ter Kuile et al., 2003), are feasible under the MFD framework developed in this paper. The assumption of no interference will likely fail but weakening Assumption 3 to allow for partial interference (Hudgens and Halloran, 2008), e.g., interference within but not between villages or sites, is plausible. Cluster randomized trial designs would allow for identification of natural definitions of treatment efficacy that are functions of the fraction of subjects who receive treatment (Athey et al., 2018). The study of erythrocytic vaccines in the MFD framework may be more challenging, as the assumption that protective hemoglobinopathies and these blood-stage vaccines don't interact is tenuous at best. With more than thirty vaccines currently under development, both pre-erythrocytic and those targeting different stages of the disease cycle, the methods described in this paper have the potential to improve the reliability of vaccine efficacy estimates for a number of forthcoming trials (Mahmoudi and Keshavarz, 2017).

Many interesting research directions related to Mendelian factorial design remain. We mention a few in closing. Developing MFD methods using only aggregate site-level data on HbAS prevalence is one such direction. This might allow for MFD-based meta-analyses of past trial results in which hemoglobinopathy data were not collected if, (1) accurate prevalence data could be collected ex post and (2) the prevalence varied sufficiently between studies. There is evidence that there are a number of other genetic traits that confer protec-

tion against malaria (Ndila et al., 2018). When there are several potential valid Mendelian factors, we may again find inspiration from MR and other IV methods. It would be beneficial to develop falsification tests of the validity of potential Mendelian factors akin to tests for over-identifying restrictions such as the Sargan-Hansen test in IV regression (Hansen, 1982). As it has been demonstrated in MR, using many ostensible Mendelian factors where some are invalid may have the potential to improve the robustness of MFD analysis (Kang et al., 2016). Finally, the application of MFD to the study of vaccines against other childhood diseases that have unspecific symptoms and no gold-standard case definition, such as pediatric tuberculosis (Beneri et al., 2016), could prove fruitful if plausible mendelian factors can be identified.

Supplemental Web Materials for *Estimating Malaria Vaccine Efficacy in the Absence of a Gold Standard Case Definition: Mendelian Factorial Design*

A Time-to-First Malaria Fever: Mendelian Factorial Design Under The Proportional Hazards Assumption

The World Health Organization (WHO) recommends that the primary endpoint in pivotal Phase III trials assessing the efficacy of malaria vaccines be the time-to-first malaria fever and that the efficacy be measured as one minus the hazard ratio returned by a Cox proportional hazard regression (Moorthy et al., 2007). In this section we use our potential outcome framework to precisely define vaccine efficacy in time-to-first malaria fever studies and show how it can be identified using MFD and estimated using data on time-to-first fever of any cause.

A.1 Parameter Identification and Estimation under Modeling Assumptions

In this section we can think about our clinical outcomes as stochastic processes indexed by t, $Y := \{Y_t\}_{t \geq 0}$. This allows us to consider quantities like the time to first fever, $T = \min\{t : Y_t = 1\}$.

If we follow the WHO recommendation above in the context of the potential outcome framework developed in §3.2, a natural quantity to study in a time-to-event analysis of vaccine efficacy is the hazard rate of malaria-attributable fevers, i.e., the instantaneous risk of developing a malaria-attributable fever at time t conditional on being free of malaria-attributable fevers up to time t. The risk set used in this hazard function includes individuals who have experienced non-malaria fevers prior to time t. This is analogous to a *subdistribution hazard* in the competing risks literature where non-malaria fevers can be viewed as a competing risk (Fine and Gray, 1999). Unfortunately, when Y^m is not observable, vaccine efficacy based on a hazard ratio using the definition of malaria-attributable fever in (1) is not

identifiable using MFD. This is due to the fact that on the hazard ratio scale, Assumption 1 (additivity) does not hold when Y^m and Y^{nm} are dependent – recall, if Y_t^{nm} equals 1 then Y_t^m must equal 0 but is otherwise free to take values in $\{0,1\}$. Only under independence (or independence conditional on X) can we additively decompose the hazard of developing a fever of any cause into the hazard of developing a malaria-attributable fever and the hazard of developing a non-malaria fever. That said, under a few additional assumptions we can still identify vaccine efficacy in terms of the hazard rate for malaria-attributable fevers in the *absence of non-malaria infections*. This hazard is analogous to a cause-specific hazard in the competing risks literature when competing risks can be considered conditionally independent (Hsu et al., 2017).

A.1.1 Malaria-attributable Fevers in the Absence of Competing Infections

If we assume that there is no combined effect of malaria and non-malaria infections on fevers (Assumption 2(iii) of Lee and Small (2018)), then we can provide an alternative decomposition of Y to the decomposition presented in (1):

$$Y(D(z,g), U(z)) = Y(D(z,g), 0) \vee Y(0, U(z)), \tag{13}$$

where $Y(D(z,g), 0)$ are malaria-attributable fevers in the absence of non-malaria infections, $Y(0, U(z))$ are non-malaria fevers in the absence of malaria infections, and $\vee$ is a pairwise maximum. For stochastic processes, $A \vee B = \{\max(A_t, B_t)\}_{t \geq 0}$. We will refer to $Y(D(z,g), 0)$ and $Y(0, U(z))$ as *isolated* malaria fevers and *isolated* non-malaria fevers. If X sufficiently captures shared risk factors for malaria and non-malaria infections, then it is plausible that $D(z,g) \perp\!\!\!\perp U(z) \mid X$ and thus $Y(D(z,g), 0) \perp\!\!\!\perp Y(0, U(z)) \mid X$ for all z, g.

One might wonder if isolated malaria fevers are the endpoint of greatest interest. Perhaps malaria-attributable fevers as defined in (1) are more representative of the real-world burden of malaria infections. Regardless, one could argue that the standard case definition of clinical malaria, e.g., $Y = 1$ and $D > 2500$ parasites per μl, is an approximation of an isolated malaria fever. This argument has two parts: (1) the fixed threshold suggests that the standard case definition does not consider the possibility of combined effects of malaria and non-malaria fevers; and (2) the standard case definition does not consider whether a non-malaria fever would have occurred had there been no malaria infection. The approximation is rough,

however, because while the definition of isolated malaria fever allows for individual-specific pyrogenic thresholds, the standard case definition does not.

A.1.2 Identifying Vaccine Efficacy

We can now define the potential time to first isolated malaria fever as $T^m(z, g) = \min\{t : Y_t(D(z, g), 0) = 1\}$ and the potential time to first isolated non-malaria fever as $T^{nm}(z, g) = \min\{t : Y_t(0, U(z)) = 1\}$. We can write the conditional hazard functions for $T^m(z, g)$ and $T^{nm}(z, g)$ as

$$\lambda_{zg}^k(t \mid X) = \lim_{\Delta \to 0_+} \mathbb{P}\big(T^k(z, g) \in [t, t + \Delta) \mid T^k(z, g) \geq t,\, X\big) / \Delta \tag{14}$$

for $k = m, nm$ and all z, g. We drop the superscript k to indicate the conditional hazard function for the a fever of any cause and define

$$T(z, g) = \min\{t : Y_t(z, g) = 1\} = \min\{T^m, T^{nm}\}.$$

Assumption 6 (Proportional Hazards). $\lambda_{zg}^k(t \mid X)$, $k = m, nm$ follow a Cox proportional hazard model with baseline hazard functions $\lambda^k(t)$, $k = m, nm$. That is,

(i) $\lambda_{zg}^m(t \mid X) = \lambda^m(t) \exp\{\log \kappa + \log(1 - \tau)z + \log(1 - \nu)g + \boldsymbol{\beta}_m^T X\}$, and

(ii) $\lambda_{zg}^{nm}(t \mid X) = \lambda^{nm}(t) \exp\{\log \phi + \log(1 - \eta)z + \boldsymbol{\beta}_{nm}^T X\}$

where $\nu > 0$.

In equation (i) above, vaccine efficacy τ is equal to one minus the hazard ratio of isolated malaria fever under vaccination versus placebo. In (ii), η can again be thought of as a "spillover efficacy" term. Note that Assumptions 2 and 5 are satisfied under these modeling assumptions.

Assumption 7 (Conditionally Independent First Fever Processes). The time to first isolated malaria fever and time to first isolated non-malaria fever are conditionally independent. That is,

$$T^m(z, g) \perp\!\!\!\perp T^{nm}(z) \mid X \quad \text{for } (z, g) \in \{0, 1\}^2$$

Because $T^m(z, g)$ depends only on $Y(D(z, g), 0)$ and $T^{nm}(z, g)$ depends only on $Y(0, U(z))$, Assumption 7 is satisfied when $Y(D(z, g), 0) \perp\!\!\!\perp Y(0, U(z)) \mid X$ for all z, g. We previously

argued that this condition is plausible when X sufficiently describes shared risk factors for malaria and non-malaria infections.

Assumption 8 (Shared Conditional Baseline Hazard Function). Time-to-first isolated malaria fever and time to first isolated non-malaria fever have a shared baseline hazard function conditional on X. That is, $\lambda(t) := \lambda^m(t) = \lambda^{nm}(t)$ for all $t \geq 0$ and $\boldsymbol{\beta} := \boldsymbol{\beta}_m = \boldsymbol{\beta}_{nm}$.

Under these additional assumptions, we can use an MFD strategy to identify τ in time-to-first fever studies.

Proposition 4 (Identification of τ). Suppose that Assumptions 6 - 8 hold. Then the hazard function for $T(z, g)$ can be written as

$$\lambda_{zg}(t \mid X) = \lambda(t) \exp\{\alpha + \omega z + \gamma g + \lambda z \times g + \boldsymbol{\beta}^T X\}. \tag{15}$$

Furthermore, under Assumptions 3 and 4, τ is identified from the observed data $\mathbf{O}$ as

$$\tau = 1 - \frac{\exp\{\omega + \gamma + \lambda\} - \exp\{\omega\}}{\exp\{\gamma\} - 1}. \tag{16}$$

Proof. The proof of Proposition 4 can be found in Appendix B.2 $\qquad\qquad\square$

Remark 2. The conditions of Proposition 4 lead to an additivity property of the conditional hazard functions that is analogous to Assumption 1. This additivity can be seen in the first equality of (26).

Remark 3. A weaker version of the randomization / "as-if" randomized condition imposed by Assumption 4 would suffice to identify τ. Namely, that Z, G are independent of all potential outcomes conditional on X.

A.1.3   Estimation and Inference via the Delta Method

The parameters in (15) can be estimated consistently via maximum partial likelihood estimation (Cox, 1975) with the `R` function `coxph` implemented in the `survival` package. These estimates $(\hat{\omega}, \hat{\gamma}, \hat{\lambda})$ are asymptotically normal with the limiting distribution

$$\sqrt{n}\left(\begin{bmatrix}\hat{\omega}\\\hat{\gamma}\\\hat{\lambda}\end{bmatrix} - \begin{bmatrix}\omega\\\gamma\\\lambda\end{bmatrix}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\begin{bmatrix}0\\0\\0\end{bmatrix}, \mathrm{I}([\omega, \gamma, \lambda^T])^{-1}\right), \tag{17}$$

where $\mathrm{I}([\omega, \gamma, \lambda^T])^{-1}$ is the inverse of the partial likelihood-based information matrix. Applying the continuous mapping theorem to (16) yields the following consistent MFD estimator of τ,

$$\hat{\tau} = 1 - \frac{\exp\{\hat{\omega} + \hat{\gamma} + \hat{\lambda}\} - \exp\{\hat{\omega}\}}{\exp\{\hat{\gamma}\} - 1}. \tag{18}$$

Applying the delta method to (17) and (18) and noting that $\hat{\mathrm{I}}([\hat{\omega}, \hat{\gamma}, \hat{\lambda}^T])^{-1}/n$, the average sample information, is consistent for $\mathrm{I}([\omega, \gamma, \lambda^T])^{-1}$ gives us an approximate distribution for $\hat{\tau}$ for large enough samples,

$$\hat{\tau} \dot{\sim} \mathcal{N}\left(\tau, \boldsymbol{\partial\hat{\tau}}^T \hat{\mathrm{I}}([\hat{\omega}, \hat{\gamma}, \hat{\lambda}^T])^{-1} \boldsymbol{\partial\hat{\tau}}\right), \tag{19}$$

where $\boldsymbol{\partial\hat{\tau}} = [\partial\hat{\tau}/\partial\hat{\omega}, \partial\hat{\tau}/\partial\hat{\gamma}, \partial\hat{\tau}/\partial\hat{\lambda}]^T$. We can use this approximate distribution to conduct inference on τ.

An analogous bounded estimator and confidence interval to those described in Proposition 3 can be constructed using the naive estimator $1 - \exp\{\hat{\omega}\}$. It can easily be shown that this estimator is consistent for a convex combination of τ and η and is thus is a consistent estimator for a lower bound of τ as long as $\eta \leq \tau$.

A.2   Causal Identification: Selection Bias and Frailty

Even if these strong modeling assumption assumptions hold, τ defined in §A.1 does not have a straightforward causal interpretation. Several authors have discussed in depth the subtleties and common misconceptions of interpreting hazard ratios as a causal quantity (Hernán, 2010; Aalen et al., 2015; Martinussen et al., 2018). We observe one of these subtleties when examining the definition of τ given in Assumption 6(i),

$$\tau = 1 - \frac{\lim_{\Delta \to 0_+} \mathbb{P}(T^m(1, g) \in [t, t + \Delta\,)|\, T^m(1, g) \geq t,\, X)}{\lim_{\Delta \to 0_+} \mathbb{P}(T^m(0, g) \in [t, t + \Delta\,)|\, T^m(0, g) \geq t,\, X)}. \tag{20}$$

These authors point out that the hazard functions in the numerator and the denominator of (20) condition on different sets of subjects for $t > t_{(1)} = \min_{ij,g,z} T_{ij}^m(g, z)$. The causal contrast is thus some combination of treatment efficacy and a selection effect. Initially, randomized assignment of Z and "as-if" randomization of G ensure that the units in each combination of trial arm and HbAS status are comparable on average. However, conditioning on the study subjects at risk of their first isolated malaria fever at time $t > t_{(1)}$, i.e.,

$T^m(z, g) \geq t$, has the potential to introduce selection bias. For instance, suppose that baseline susceptibility to malaria is heterogeneous even after conditioning on X. If Z provides protection against developing an isolated malaria fever, we might find that the children with high susceptibility to malaria in the control arm are more likely to develop fevers than similar children in the treatment arm. Consequently, as time passes, the children at risk in the control arm will be less susceptible on average to fever than those in the treatment arm, absent treatment. The comparability of the subjects in each combination of $z, g \in \{0, 1\}^2$ ensured by Assumption 4 at the beginning of the study is not guaranteed as the follow-up progresses.

When malaria-attributable fever is rare or when the follow-up time is short, the selection effect may be less consequential (Aalen et al., 2015). If X does not sufficiently capture heterogenous susceptibility to isolated malaria fever, modeling subject-level frailty can further alleviate the selection bias of estimates of τ (Wienke, 2010). Frailty can be introduced as a multiplicative random effect,

$$\lambda_{zg}^m(t \mid X) = \lambda(t)W \exp\{\log \alpha + \log(1 - \tau)z + \log(1 - \nu)g + \boldsymbol{\beta}^T X\},$$

$$\lambda_{zg}^{nm}(t \mid X) = \lambda(t)W \exp\{\log \phi + \log(1 - \eta)z + \boldsymbol{\beta}^T X\} \tag{21}$$

where W is a random, subject-level frailty shared by both malaria and non-malaria fever hazard functions. Proposition 4 extends immediately to the frailty model implied by (21) where, in addition to X, the hazard ratio is conditional on W. Estimation can be carried out via the Expectation-Maximization algorithm (Klein, 1992) or penalized partial maximum likelihood methods (Therneau et al., 2003).

Alternatively, frailty models can be used to conduct a sensitivity analysis of Cox regression-based estimates of τ to selection bias (Stensrud et al., 2017). In general, however, modeling frailty usually requires parametric assumptions about W, for example, that it comes from the family of power variance function distributions (Wienke, 2010).

Although τ itself has a subtle and potentially awkward causal interpretation, simple functions of τ have been shown to have more natural causal interpretations. For example, $1/(2 - \tau)$ is the *probabilistic index*, which is defined as the probability that $T^m(1, g)$ for one individual is longer than $T^m(0, g)$ for another individual with comparable baseline covariates (and frailty) (De Neve and Gerds, 2019).

Importantly, all the causal interpretations discussed in this section require the correct identification of the parameter τ established in Proposition 4.

B   Technical Appendices

B.1   Proof of Proposition 1

In this Appendix we give a proof of the general identification result in Proposition 1.

Proof of Proposition 1. By consistency and Assumption 3 we have that

$$\mathbb{E}_{P_j}[f(Y)|\,Z=z,G=g,X] = \mathbb{E}_{P_j}[f(Y(z,g))|\,Z=z,G=g,X] \tag{22}$$

for all z,g. Assumption 4 ensures that $P_{X|Z,G,j} = P_{X,j}$ and $P_{Y(z,g),j|Z,G} = P_{Y(z,g),j}$ where the subscript j indicates that these are conditional distributions from a factorization of the site-specific P_j. Hence,

$$
\begin{aligned}
\mathbb{E}_{X,j}\big[\mathbb{E}_{P_j}[f\{Y(z,g)\}|\,X,Z=z,G=g]\big] &= \mathbb{E}_{X|Z=z,G=g,j}\big[\mathbb{E}_{P_j}[f\{Y(z,g)\}|\,X,Z=z,G=g]\big]\\
&= \mathbb{E}_{P_j}[f\{Y(z,g)\}|\,Z=z,G=g]\\
&= \mathbb{E}_{P_j}[f\{Y(z,g)\}]\\
&= \mu_{zg}(P_j)\,.
\end{aligned}
\tag{23}
$$

Noting that $J^{-1}\sum_{j=1}^{J}\mu_{zg}(P_j) = \mu_{zg}(P)$ and applying Assumption 1, the right hand side of (2) becomes

$$1 - \frac{(\mu_{11}^m(P) - \mu_{10}^m(P)) + (\mu_{11}^{nm}(P) - \mu_{10}^{nm}(P))}{(\mu_{01}^m(P) - \mu_{00}^m(P)) + (\mu_{01}^{nm}(P) - \mu_{00}^{nm}(P))}\,. \tag{24}$$

Because G is a valid Mendelian factor (Assumption 5(ii)) we have that $\mu_{z1}^{nm}(P) - \mu_{z0}^{nm}(P) = 0$ for $z = 0,1$, further simplifying (24) to

$$1 - \frac{(\mu_{11}^m(P) - \mu_{10}^m(P))}{(\mu_{01}^m(P) - \mu_{00}^m(P))}\,.$$

Finally, we have that

$$
\begin{aligned}
1 - \frac{(\mu_{11}^m(P) - \mu_{10}^m(P))}{(\mu_{01}^m(P) - \mu_{00}^m(P))} &= 1 - \frac{-\nu\mu_{10}^m(P)}{-\nu\mu_{00}^m(P)} \qquad \text{by Assumption 2 and Definition 1(ii)}\\
&= 1 - \frac{\mu_{10}^m(P)}{\mu_{00}^m(P)}\\
&= \tau \qquad\qquad\qquad \text{by Definition 1(i)}\,.
\end{aligned}
$$

Assumption 5(i) ensures that the right hand side of the first equality is well-defined. $\qquad \square$

B.2 Proof of Proposition 3

We give a short proof of the asymptotic unbiasedness of $\hat{\tau}_{bnd}$ and the asymptotical validity of $\mathrm{CI}_{bnd,\alpha}$ as a $1 - \alpha$ confidence interval for τ.

Proof of Proposition 3. We give a proof for when the sample sizes are balanced across sites. When this is the case, we skip step 2 of Algorithm 1 and use $\hat{\mu}_0$ to construct $\hat{\tau}_0$. Treating G as an element of X, the consistency of $\hat{\tau}_0$ for $s\tau + (1-s)\eta$ and its asymptotic linearity follows almost immediately from Theorem 1 of Rosenblum and van der Laan (2010). When $\eta < \tau$, the theorem implies that $\hat{\tau}_0$ is consistent for some $\tau' < \tau$. Because $\tau \le 1$ by definition, the continuous mapping theorem implies that $\hat{\tau}_{bnd}$ is consistent for $\min[1, \max\{\tau', \tau\}] = \tau$. The lower bound of $\mathrm{CI}_{bnd,\alpha}$ is constructed by taking the intersection of a lower one-sided confidence interval for τ with coverage $1 - \alpha/2 + \alpha_0$ and a lower one-sided confidence interval for τ' with coverage $1 - \alpha_0$. Because $\tau' < \tau$, this second confidence interval is also asymptotically valid for τ. If α_0 is chosen a priori, then the intersection is an asymptotically valid $1 - \alpha/2$ lower one-sided interval for τ (Neuwald and Green, 1994). The upper confidence bound is constructed by taking the intersection of a $1 - \alpha/2$ upper one-sided confidence interval for τ and $(-\infty, 1]$. Since $\tau \le 1$ the resulting interval is an asymptotically valid $1 - \alpha/2$ upper confidence-interval for τ. The intersection of the resultant upper and lower $1 - \alpha/2$ one-sided confidence intervals yields an asymptotically valid two-sided confidence interval with coverage $1 - \alpha$. □

B.3 Proof of Proposition 4

In this Appendix we provide a proof of the parameter identification result for the hazard ratio under the proportional hazards assumption.

Proof of Proposition 4. Let $\Lambda_{zg}(t \mid X) = \int_{-\infty}^{t} \lambda_{zg}(t \mid X)\, dt$ be the cumulative hazard function for Y. Define $\Lambda_{zg}^k(t \mid X)$, $k = m, nm$ similarly. The survival function $S_{zg}(t \mid X = \mathbb{P}(T(z, g) > t \mid X)$ can be expressed as $S_{zg}(t \mid X) = \exp\{-\Lambda_{zg}(t \mid X)\}$ and the probability

density of Y as $f_{zg}(t \mid X) = \lambda_{zg}(t \mid X)S_{zg}(t \mid X)$. We can express $f_{zg}(t \mid X)$ as

$$
\begin{aligned}
f_{zg}(t \mid X) &= \frac{\partial}{\partial t} F_{zg}(t \mid X) \\
&= \frac{\partial}{\partial t} \left\{ 1 - S_{zg}^{nm}(t \mid X) S_{zg}^{m}(t \mid X) \right\} \quad \text{by Assumption 7} \\
&= f_{zg}^{m}(t \mid X) S_{zg}^{nm}(t \mid X) + f_{zg}^{nm}(t \mid X) S_{zg}^{m}(t \mid X) \\
&= \left\{ \lambda_{zg}^{m}(t \mid X) + \lambda_{zg}^{nm}(t \mid X) \right\} S_{zg}^{nm}(t \mid X) S_{zg}^{m}(t \mid X) \\
&= \left\{ \lambda_{zg}^{m}(t \mid X) + \lambda_{zg}^{nm}(t \mid X) \right\} \times \exp \left\{ - \int_{-\infty}^{t} \lambda_{zg}^{m}(t \mid X) + \lambda_{zg}^{nm}(t \mid X)\, dt \right\}.
\end{aligned}
\tag{25}
$$

From (25) and Assumptions 6 and 8, we have that the hazard function for Y is

$$
\begin{aligned}
\lambda_{zg}(t \mid X) &= \lambda_{zg}^{m}(t \mid X) + \lambda_{zg}^{nm}(t \mid X) \\
&= \lambda(t) \exp\{\boldsymbol{\beta}^{T} X\} \left(\kappa(1-\tau)^{z}(1-\nu)^{g} + (1-\eta)^{z} \right) \\
&= \lambda(t) \exp\{\alpha + \omega z + \gamma g + \lambda z \times g + \boldsymbol{\beta}^{T} X\},
\end{aligned}
\tag{26}
$$

Proving the first part of the proposition. To prove the second part of the proposition we first observe that the last equality above in (26) is simply a reparameterization – when z and g are binary, $\kappa(1-\tau)^{z}(1-\nu)^{g} + \phi(1-\eta)^{z}$ and $\exp\{\alpha + \omega z + \gamma g + \lambda z \times g\}$ can each take four distinct values. The reparameterization yields a system of four equations,

$$
\exp\{\alpha\} = \kappa + \phi
\tag{27}
$$

$$
\exp\{\alpha + \omega\} = \kappa(1-\tau) + \phi(1-\eta)
\tag{28}
$$

$$
\exp\{\alpha + \gamma\} = \kappa(1-\nu)
\tag{29}
$$

$$
\exp\{\alpha + \omega + \gamma + \lambda\} = \kappa(1-\tau)(1-\nu) + \phi(1-\eta).
\tag{30}
$$

Subtracting (28) from (30) and (27) from (29), then dividing the former by the latter yields

$$
1 - \tau = \frac{\exp\{\omega + \gamma + \lambda\} - \exp\{\omega\}}{\exp\{\gamma\} - 1}.
\tag{31}
$$

Under Assumptions 3 and 4 we have that $\lambda_{zg}(t \mid X)$ can be identified by the observed data **O**:

$$
\begin{aligned}
\lambda_{zg}(t \mid X) &= \lim_{\Delta \to 0_{+}} \mathbb{P}(T(z,g) \in [t, t+\Delta] \mid T(z,g) \geq t,\, X)/\Delta \\
&= \lim_{\Delta \to 0_{+}} \mathbb{P}(T \in [t, t+\Delta] \mid T \geq t,\, X,\, Z = z,\, G = g)/\Delta.
\end{aligned}
\tag{32}
$$

Applying (26), we have that $\alpha, \omega, \gamma, \lambda$ are identifiable and thus τ can be identified by the observed data using (31). $\qquad \square$

B.4 Proof (Sketch) of Proposition 2

In this section we sketch a proof of Proposition 2 by showing that $\mu_{zg}(P_n)$ is consistent for $\mu_{zg}(P)$ and asymptotically linear, that is,

$$\mu_{zg}(P_n) - \mu_{zg}(P) = (\mathbb{P}_n - \mathbb{P})\varphi^*_{zg}(P_\infty) + o_P\left(1/\sqrt{N}\right). \tag{33}$$

for all z, g. Theorem 1 of Rosenblum and van der Laan (2010) checks a number of technical conditions to verify that Theorem 1 of van der Laan and Rubin (2006) can be applied. A particularly important condition is that the $\mu_{zg}(P)$ are linear in P. The implication of the parameter μ_{zg} being linear is that we can write

$$\mu_{zg}(P_n) - \mu_{zg}(P) = (\mathbb{P}_n - \mathbb{P})\varphi^*_{zg}(P_n). \tag{34}$$

We demonstrate that this equality holds under our class of models.

If $\varphi^*_{zg}(P_n)$ is Donsker then by the second part of Theorem 1 of van der Laan and Rubin (2006), we have that $\mu_{zg}(P_n)$ is $\sqrt{n}$-consistent for $\mu_{zg}(P)$. Then, if $\mathbb{P}\{\varphi^*_{zg}(P_n) - \varphi^*_{zg}(P_\infty)\}^2 \to 0$ in probability as $n \to \infty$ the third part of Theorem 1 of van der Laan and Rubin (2006) implies that $\mu_{zg}(P_n)$ is asymptotically linear with form (33). Consistency of $\mu_{zg}(P_n)$ and Assumptions 1 - 5 imply that the plugin estimator $\hat{\tau}$ is consistent for τ. Finally, we apply the delta method to derive the asymptotic variance σ^2 of $\sqrt{n}(\hat{\tau} - \tau)$. It follows from the last part of Theorem 1 of van der Laan and Rubin (2006) that σ^2 achieves the semiparametric efficiency bound if the working model for the conditional expectation of $f(Y)$ is correctly specified. In what follows, when taking expectations of functionals of P_n we treat these functionals as fixed.

Verifying (34):

Given the choice of the terms included in the linear part of the GLM estimate $\hat{\mu}_{j,1}$ in Algorithm 1, the score equations that $\hat{\mu}_{j,1}$ solve imply that

$$\mathbb{P}_n\left\{\frac{\mathbb{1}(Z = z)\mathbb{1}(G = g)}{p(Z = z)p_{j,n}(G = g)}\left(f(Y) - \hat{\mu}_{j,1}(z, g, X)\right)\right\} = 0 \quad \text{for all} \quad z, g \in \{0, 1\}^2. \tag{35}$$

The choice of the weighting $\mathtt{w} = n/(J \cdot I_j)$ when estimating $\hat{\mu}_1$ is required since $\mathbb{P}_n$ does not equally weight observations unless the size of each site is the same. Also note, that by

definition, $\mathbb{P}_n \hat{\mu}_{j,1}(z,g,X) = \mu_{zg}(P_n)$. Taken together, with (35) we have that $\mathbb{P}_n \varphi_{zg}(P_n) = 0$ for all $z, g \in \{0,1\}^2$. By iterated expectations, we have that

$$\mathbb{P}\varphi_{zg}^*(P_n) = \mathbb{P}\varphi_{zg}(P_n) - \mathbb{P}\{\mathbb{E}_P[\varphi_{zg}(P_n) \mid G]\} = \mathbb{P}\varphi_{zg}(P_n) - \mathbb{P}\varphi_{zg}(P_n) = 0. \qquad (36)$$

All that remains to be shown is that $-\mathbb{P}_n \{\mathbb{E}_P[\varphi_{zg}(P_n) \mid G]\} = \mu_{zg}(P_n) - \mu_{zg}(P)$, or equivalently, that $-\mathbb{P}_n \{\mathbb{E}_{P_j}[\varphi_{zg}(P_{j,n}) \mid G]\} = \mu_{zg}(P_{j,n}) - \mu_{zg}(P_j)$. We start by evaluating $\mathbb{E}_{P_j}[\varphi_{zg}(P_{j,n}) \mid G]$:

$$\mathbb{E}_{P_j}[\varphi_{zg}(P_{j,n}) \mid G] = \mathbb{E}_{P_j}\left[\frac{\mathbb{1}(Z=z)\mathbb{1}(G=g)\{f(Y) - \hat{\mu}_{j,1}(z,g,X)\}}{p_{j,n}(G=g)p(Z=z)} \,\Big|\, G\right]$$
$$+ \mathbb{E}_{P_j}[\hat{\mu}_{j,1}(z,g,X) - \mu_{zg}(P_{j,n}) \mid G]$$
$$= \frac{\mathbb{1}(G=g)}{p_{j,n}(G=g)}\left(\mu_{zg}(P_j) - \mathbb{E}_{P_j}[\hat{\mu}_{j,1}(z,g,X) \mid G]\right)$$
$$+ \mathbb{E}_{P_j}[\hat{\mu}_{j,1}(z,g,X) \mid G] - \mu_{zg}(P_{j,n}). \qquad (37)$$

The second equality follows from an application of iterated expectations, Assumption 4, and the fact that $\mu_{zg}(P_{j,n})$ and $\hat{\mu}_{j,1}$ are treated as fixed when taking expectations. Taking the empirical expectation $\mathbb{P}_n$ we get

$$\mathbb{P}_n\{\mathbb{E}_{P_j}[\varphi_{zg}(P_{j,n}) \mid G]\} = \mathbb{P}_n\left\{\frac{\mathbb{1}(G=g)}{p_{j,n}(G=g)}\left(\mu_{zg}(P_j) - \mathbb{E}_{P_j}[\hat{\mu}_{j,1}(z,g,X) \mid G]\right)\right\}$$
$$+ \mathbb{P}_n\{\mathbb{E}_{P_j}[\hat{\mu}_1(z,g,X) \mid G] - \mu_{zg}(P_{j,n})\}$$
$$= \mu_{zg}(P_j) - \mathbb{E}_{P_j}[\hat{\mu}_{j,1}(z,g,X) \mid G=g]$$
$$+ \mathbb{P}_n\{\mathbb{E}_P[\hat{\mu}_{j,1}(z,g,X) \mid G]\} - \mu_{zg}(P_{j,n})$$
$$= \mu_{zg}(P_j) - \mathbb{E}_{P_j}[\hat{\mu}_{j,1}(z,g,X)] + \mathbb{E}_{P_j}[\hat{\mu}_{j,1}(z,g,X)] - \mu_{zg}(P_{j,n})$$
$$= -(\mu_{zg}(P_{j,n}) - \mu_{zg}(P_j)). \qquad (38)$$

The second to last equality comes from the fact that $X \perp\!\!\!\perp G$ by Assumption 4. This is our desired result and (34) now follows.

Checking that $\varphi_{zg}^*(P_n)$ is Donsker:

This is analogous to condition (iv) in the proof of Theorem 1 in Rosenblum and van der Laan (2010). We make the same assumptions of boundedness on $\boldsymbol{\beta}$ and on the terms in the linear part of the generalized linear models in steps 1 and 2 of Algorithm 1. Under the

additional assumption that $0 < \delta \leq p_j(G = g)$ for all $g = 0, 1$ and $j = 1, \ldots, J$, a nearly identical verification that $\varphi_{zg}^*(P_n)$ is Donsker follows. Hence, $\mu_{zg}(P_n)$ are $\sqrt{n}$-consistent for all $z, g \in \{0, 1\}1^2$.

Verifying that $\mathbb{P}\{\varphi_{zg}^*(P_n) - \varphi_{zg}^*(P_\infty)\}^2 \xrightarrow{P} 0$:

The steps to verify that $\mathbb{P}\{\varphi_{zg}^*(P_n) - \varphi_{zg}^*(P_\infty)\}^2 \to 0$ in probability are very similar to the verification of condition (v) in Rosenblum and van der Laan (2010). There are a few extra steps required to deal with the fact that we are estimating $p_j(G = g)$ with $p_{j,n}(G = g)$. We first note that we can write

$$
\begin{aligned}
\mathbb{P}\{\varphi_{zg}^*(P_n) - \varphi_{zg}^*(P_\infty)\}^2 &= \mathbb{P}[\{\varphi_{zg}(P_n) - \varphi_{zg}(P_\infty)\} - \mathbb{E}_P[\varphi_{zg}(P_\infty) - \varphi_{zg}(P_n) \mid G]]^2 \\
&\leq 2\mathbb{P}\{\varphi_{zg}(P_n) - \varphi_{zg}(P_\infty)\}^2 + 2\mathbb{P}\{\mathbb{E}_P[\varphi_{zg}(P_\infty) - \varphi_{zg}(P_n) \mid G]\}^2 \\
&\leq 4\mathbb{P}\{\varphi_{zg}(P_n) - \varphi_{zg}(P_\infty)\}^2 \\
&\leq 4J^{-1}\sum_{i=1}^{J}\mathbb{P}\{\varphi_{zg}(P_{j,n}) - \varphi_{zg}(P_{j,\infty})\}^2.
\end{aligned}
\tag{39}
$$

The second to last line comes from applying Jensen's inequality to the inner expectation of the second term followed by an application of iterated expectations. The last line comes from another application of Jensen's inequality. We need to show that $\mathbb{P}\{\varphi_{zg}(P_{j,n}) - \varphi_{zg}(P_{j,\infty})\}^2$ converges to 0 in probability. In what follows we distinguish the working model using the fitted parameters $\boldsymbol{\beta}_n$ and the model using the unique maximizer of the expected log-likelihood $\boldsymbol{\beta}$ as $\hat{\mu}_{j,1,n}$ and $\hat{\mu}_{j,1,\infty}$, respectively. For brevity, we also let $\mathbb{1}_{zg} = \mathbb{1}(Z = z)\mathbb{1}(G = g)$,

$p(z) = p(Z = z)$, $p_j(g) = p_j(G = g)$, and $p_{j,n}(g) = p_{j,n}(G = g)$. We can write

$$
\begin{aligned}
\mathbb{P}&\{\varphi_{zg}(P_{j,n}) - \varphi_{zg}(P_{j,\infty})\}^2 \\
&= \mathbb{P}\left\{\frac{\mathbb{1}_{zg}}{p(z)p_{j,n}(g)}\{f(Y) - \hat{\mu}_{j,1,n}(z,g,X)\} + \hat{\mu}_{j,1,n}(z,g,X) - \mu_{zg}(P_{j,n})\right. \\
&\qquad\left. - \frac{\mathbb{1}_{zg}}{p(z)p_j(g)}\{f(Y) - \hat{\mu}_{j,1,\infty}(z,g,S)\} - \hat{\mu}_{j,1,\infty}(z,g,X) + \mu_{zg}(P_j)\right\}^2 \\
&\le \frac{4}{p(z)^2}\mathbb{P}\left\{f(Y)^2\right\}\left(\frac{1}{p_{j,n}(g)} - \frac{1}{p_j(g)}\right)^2 + \frac{4}{p(z)^2}\mathbb{P}\left\{\frac{\hat{\mu}_{j,1,\infty}(z,g,X)}{p_j(g)} - \frac{\hat{\mu}_{j,1,n}(z,g,X)}{p_{j,n}(g)}\right\}^2 \\
&\qquad + 4\mathbb{P}\left\{\hat{\mu}_{j,1,n}(z,g,X) - \hat{\mu}_{j,1,\infty}(z,g,X)\right\}^2 + 4\mathbb{P}\left\{\mu_{zg}(P_j) - \mu_{zg}(P_{j,n})\right\}^2 \\
&\le C_1\left(\frac{1}{p_{j,n}(g)} - \frac{1}{p_j(g)}\right)^2 + \frac{4}{p(z)^2}\mathbb{P}\left\{\frac{\hat{\mu}_{j,1,\infty}(z,g,X)}{p_j(g)} - \frac{\hat{\mu}_{j,1,n}(z,g,X)}{p_{j,n}(g)}\right\}^2 \\
&\qquad + C_2\|\boldsymbol{\beta}_0 - \boldsymbol{\beta}\|^2 + 4\{\mu_{zg}(P_j) - \mu_{zg}(P_{j,n})\}^2.
\end{aligned}
\tag{40}
$$

The first equality follows immediately from definitions. The first *inequality* follows from rearranging terms and noting that $(x_1 + \cdots + x_k)^2 \le k(x_1^2 + \cdots + x_k^2)$ by Jensen's inequality. The first term after the second inequality comes from the boundedness of Y. Since $p_{j,n}(g) \xrightarrow{P} p_j(g)$ and we have assumed that $p_j(g)$ are bounded away from zero, the continuous mapping theorem implies that this term converges to 0 in probability. The third term comes from the fact that our working model has uniformly bounded first derivatives. This is due to the boundedness assumptions on X and the terms of the linear part of our working model. The conditions given in Proposition 2, are sufficient for $\boldsymbol{\beta}_n$ to converge to $\boldsymbol{\beta}$ in probability (Rosenblum and van der Laan (2009), Appendix D). Consequently, the third term also converges to 0 in probability. The fourth term disappears in probability since we proved earlier that $\mu_{zg}(P_n)$ is consistent for $\mu_{zg}(P)$. All that is left to handle is the second term. A

little bit of rearranging gives us

$$\frac{4}{p(z)^2}\mathbb{P}\left\{\frac{\hat{\mu}_{j,1,\infty}(z,g,X)}{p_j(g)} - \frac{\hat{\mu}_{j,1,n}(z,g,X)}{p_{j,n}(g)}\right\}^2$$

$$= \frac{4}{p(z)^2}\mathbb{P}\left\{\frac{\hat{\mu}_{j,1,\infty}(z,g,X) - \hat{\mu}_{j,1,n}(z,g,X,S)}{p_j(g)} + \hat{\mu}_{j,1,n}(z,g,X)\left(\frac{1}{p_{j,n}(g)} - \frac{1}{p_j(g)}\right)\right\}^2$$

$$\leq \frac{8}{\{p(z)p_j(g)\}^2}\mathbb{P}\left\{\hat{\mu}_{j,1,\infty}(z,g,X) - \hat{\mu}_{j,1,n}(z,g,X)\right\}^2$$

$$+ \frac{8}{p(z)^2}\mathbb{P}\left\{\hat{\mu}_{j,1,n}(z,g,X)^2\right\}\left(\frac{1}{p_{j,n}(g)} - \frac{1}{p_j(g)}\right)^2$$

$$= C_3\|\boldsymbol{\beta}_0 - \boldsymbol{\beta}\|^2 + C_4\left(\frac{1}{p_{j,n}(g)} - \frac{1}{p_j(g)}\right)^2. \tag{41}$$

The first inequality follows again from Jenson's inequality. The first term in the last line follows from the same arguments made earlier and the second term follows from the fact that $\mathbb{P}\{\hat{\mu}_{j,1,n}(z,g,X)^2\}$ is bounded. As we've already demonstrated, these two terms vanish in probability. Combining (39), (40), and (41) gives us that $\mathbb{P}\{\varphi_{zg}^*(P_n) - \varphi_{zg}^*(P_\infty)\}^2 \xrightarrow{P} 0$. (35) follows immediately and an application of Proposition 1 implies that $\hat{\tau}$ is also asymptotically linear. All that remains is to compute its asymptotic variance.

Asymptotic Variance:

To compute the asymptotic variance of $\hat{\tau}$ we begin with a taylor expansion around τ:

$$\hat{\tau} = 1 - \frac{\mu_1(P_n)}{\mu_0(P_n)}$$

$$= 1 - \frac{\mu_1(P)}{\mu_0(P)} - \frac{1}{\mu_0(P)}\{\mu_1(P_n) - \mu_1(P)\}$$

$$+ \frac{\mu_1(P)}{\mu_0(P)^2}\{\mu_0(P_n) - \mu_0(P)\} + o_p\left(1/\sqrt{n}\right)$$

$$= \tau - \frac{1}{\mu_0(P)}(\mathbb{P}_n - \mathbb{P})\varphi_1^*(P_\infty)(O) + \frac{\mu_1(P)}{\mu_0(P)^2}(\mathbb{P}_n - \mathbb{P})\varphi_0^*(P_\infty)(O) + o_p(1/\sqrt{n}). \tag{42}$$

The $o_p(1/\sqrt{n})$ comes from the second order remainder term of the expansion. Using the expansion in (42) along with the fact that $\varphi_z^*(P_\infty)(O) - \mathbb{P}\varphi_z^*(P_\infty)(O)$ are mean zero i.i.d. random variables we can write the scaled asymptotic variance of $\hat{\tau}$, $n \cdot \text{var}(\hat{\tau})$, as

$$\tilde{C} \cdot \mathbb{P}\left[\frac{\mu_1(P)}{\mu_0(P)^2}(\varphi_0^*(P_\infty)(O) - \mathbb{P}\varphi_0^*(P_\infty)(O)\} - \frac{1}{\mu_0(P)}\{\varphi_1^*(P_\infty)(O) - \mathbb{P}\varphi_1^*(P_\infty)(O)\}\right]^2 \tag{43}$$

where $\tilde{C} = (nJ^{-2})\sum_{j=1}^{J}I_j^{-1} \geq 1$ is a correction factor to account for the fact that sample

sizes might differ across study site due to design rather than chance. When the overall study population is balanced across sites, $\tilde{C} = 1$.

B.5 Additional Simulation Study Details for TMLE-based Estimators

Below are the simulation settings for the individual-specific vaccine and protective efficacies (τ_i and ν_i), the baseline covariates and how they are transformed when they enter the condition mean of the distribution of fever counts (X_i, $\tilde{X}_i^m$, and $\tilde{X}_i^{nm}$), and unobserved heterogeneity in the distribution of fever counts between individuals (ϵ_i^m and ϵ_i^{nm}).

Generative Distributions:

$$(1 - \nu_i) \sim (1 - \nu) \exp\{-0.05^2/2\} \cdot \texttt{Lognormal}(0, 0.05^2)$$

$$(1 - \tau_i) \sim (1 - \tau) \exp\{-0.05^2/2\} \cdot \texttt{Lognormal}(0, 0.05^2),$$

$$\tilde{X}_i^m \sim \exp\{0.05 X_i - 0.05^2/2\},$$

$$\tilde{X}_i^{nm} \sim \exp\{0.075 X_i - 0.075^2/2\},$$

$$X_i \sim \texttt{Normal}(0, 1),$$

$$\epsilon_i^m \sim \exp\{-0.05^2/2\} \cdot \texttt{Lognormal}(0, 0.05^2),$$

$$\epsilon_i^{nm} \sim \exp\{-0.05^2/2\} \cdot \texttt{Lognormal}(0, 0.05^2).$$

We assume that the spillover efficacy η is equal to zero. Based on empirical evidence and the negative dependence between malaria-attributable fevers and non-malaria fevers as defined in (1), we simulate the number of malaria-attributable and non-malaria fevers from negatively dependent negative binomial distributions (Olotu et al., 2013). We use a overdispersion parameter of $r = 10$

Count of malaria-attributable Fevers per child-year:

On the margin, the count of malaria-attributable fevers per child-year follows a negative binomial distribution with mean $\mu_{m,i}$ and variance $\sigma_{m,i}^2$,

$$\mathbf{1}^T Y_i^m(z, g) \sim \texttt{NB}(\mu_{m,i}, \sigma_{m,i}^2) \tag{44}$$

48

where $\mu_{m,i} = \kappa \cdot (1 - \nu_i)^g \cdot (1 - \tau_i)^z \cdot \tilde{X}_i^m \cdot \epsilon_i^m$ and $\sigma_{m,i}^2 = \mu_{m,i}^2/r + \mu_{m,i}$.

Count of non-malaria fevers per child-year:

On the margin, the count of non-malaria fevers per child-year follows a negative binomial distribution with mean $\mu_{nm,i}$ and variance $\sigma_{nm,i}^2$,

$$\mathbf{1}^T Y_i^{nm}(z, g) \sim \text{NB}(\mu_{nm,i}, \sigma_{nm,i}^2) \tag{45}$$

where $\mu_{nm,i} = \phi \cdot \tilde{X}_i^{nm} \cdot \epsilon_i^{nm}$ and $\sigma_{nm,i}^2 = \mu_{nm,i}^2/r + \mu_{nm,i}$.

Count of fevers of any-cause per child-year:

The joint distribution of the counts of non-malaria and malaria-attributable fevers per child-year are simulated using a Gaussian copula with a negative dependence parameter $\rho = -0.1$ (Genest and Neslehová, 2007).

Specificity and Mendelian Gene Prevalence Settings:

We calibrated κ and ϕ to achieve different levels of case specificity ($s = 0.5$ and 0.8). For example, for a case specificity of 0.8 we set κ and ϕ so that the expected number of malaria-attributable fevers is 80% of the expected number of total fevers of any-cause. The prevalence of the Mendelian gene was set to 20% based on existing estimates of HbAS prevalence in sub-Saharan Africa (Ter Kuile et al., 2003; Elguero et al., 2015).

Initial and Updated Working Model Estimates:

For both the MFD and naive estimators, we used the R function `glm.fit` from the package `stats` to fit the initial estimator $\hat{\mu}_0$ in step 1 of Algorithm 1 using Poisson regression. The regression included main terms for Z, G, and X as well as all interactions.

References

Aalen, O. O., R. J. Cook, and K. Røysland (2015). Does cox analysis of a randomized survival study yield a causal treatment effect? *Lifetime data analysis 21*(4), 579–593.

Athey, S., D. Eckles, and G. W. Imbens (2018). Exact p-values for network interference. *Journal of the American Statistical Association 113* (521), 230–240.

Bejon, P., J. A. Berkley, T. Mwangi, E. Ogada, I. Mwangi, K. Maitland, T. Williams, J. A. G. Scott, M. English, B. S. Lowe, et al. (2007). Defining childhood severe falciparum malaria for intervention studies. *PLoS medicine 4* (8), e251.

Bejon, P., M. T. White, A. Olotu, K. Bojang, J. P. Lusingu, N. Salim, N. N. Otsyula, S. T. Agnandji, K. P. Asante, S. Owusu-Agyei, et al. (2013). Efficacy of rts, s malaria vaccines: individual-participant pooled analysis of phase 2 data. *The Lancet infectious diseases 13* (4), 319–327.

Beneri, C., L. Aaron, S. Kim, P. Jean-Philippe, S. Madhi, A. Violari, M. F. Cotton, C. Mitchell, S. Nachman, et al. (2016). Understanding nih clinical case definitions for pediatric intrathoracic tb by applying them to a clinical trial. *The International Journal of Tuberculosis and Lung Disease 20* (1), 93–100.

Berger, R. L. and D. D. Boos (1994). P values maximized over a confidence set for the nuisance parameter. *Journal of the American Statistical Association 89* (427), 1012–1016.

Burgess, S., D. S. Small, and S. G. Thompson (2017). A review of instrumental variable estimators for mendelian randomization. *Statistical methods in medical research 26* (5), 2333–2355.

Cox, D. R. (1975). Partial likelihood. *Biometrika 62* (2), 269–276.

Davey Smith, G. and S. Ebrahim (2003, 02). ?Mendelian randomization?: can genetic epidemiology contribute to understanding environmental determinants of disease?*. *International Journal of Epidemiology 32* (1), 1–22.

De Neve, J. and T. A. Gerds (2019, Jan). On the interpretation of the hazard ratio in cox regression. *Biometrical Journal*.

Elguero, E., L. M. Délicat-Loembet, V. Rougeron, C. Arnathau, B. Roche, P. Becquart, J.-P. Gonzalez, D. Nkoghe, L. Sica, E. M. Leroy, et al. (2015). Malaria continues to select for

sickle cell trait in central africa. *Proceedings of the National Academy of Sciences 112*(22), 7051–7054.

Ferreira, A., I. Marguti, I. Bechmann, V. Jeney, Â. Chora, N. R. Palha, S. Rebelo, A. Henri, Y. Beuzard, and M. P. Soares (2011). Sickle hemoglobin confers tolerance to plasmodium infection. *Cell 145*(3), 398–409.

Fine, J. P. and R. J. Gray (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association 94*(446), 496–509.

Gari, T., E. Loha, W. Deressa, T. Solomon, and B. Lindtjørn (2018, Jan). Malaria increased the risk of stunting and wasting among young children in ethiopia: Results of a cohort study. *PLOS ONE 13*(1), e0190983.

Genest, C. and J. Neslehová (2007). A primer on copulas for count data. *ASTIN Bulletin: The Journal of the IAA 37*(2), 475–515.

Gong, L., S. Parikh, P. J. Rosenthal, and B. Greenhouse (2013). Biochemical and immunological mechanisms by which sickle cell trait protects against malaria. *Malaria Journal 12*(1), 317.

Gravenor, M. and D. Kwiatkowski (1998). An analysis of the temperature effects of fever on the intra-host population dynamics of plasmodium falciparum. *Parasitology 117*(2), 97–105.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Hernán, M. A. (2010, Jan). The hazards of hazard ratios. *Epidemiology 21*(1), 13–15.

Hsu, J. Y., J. A. Roy, D. Xie, W. Yang, H. Shou, A. H. Anderson, J. R. Landis, C. Jepson, M. Wolf, T. Isakova, et al. (2017). Statistical methods for cohort studies of ckd: survival analysis in the setting of competing risks. *Clinical Journal of the American Society of Nephrology 12*(7), 1181–1189.

Hudgens, M. G. and M. E. Halloran (2008). Toward causal inference with interference. *Journal of the American Statistical Association 103*(482), 832–842.

Imbens, G. W. and P. R. Rosenbaum (2005). Robust, accurate confidence intervals with a weak instrument: Quarter of birth and education. *Journal of the Royal Statistical Society. Series A (Statistics in Society) 168*(1), 109–126.

Kang, H., B. Kreuels, O. Adjei, R. Krumkamp, J. May, and D. S. Small (2013, Aug). The causal effect of malaria on stunting: a mendelian randomization and matching approach. *International Journal of Epidemiology 42*(5), 1390–1398.

Kang, H., A. Zhang, T. T. Cai, and D. S. Small (2016). Instrumental variables estimation with some invalid instruments and its application to mendelian randomization. *Journal of the American Statistical Association 111*(513), 132–144.

Kennedy, E. H. (2016). Semiparametric theory and empirical processes in causal inference. In *Statistical Causal Inferences and Their Applications in Public Health Research*, pp. 141–167. Springer.

Klein, J. P. (1992). Semiparametric estimation of random effects using the cox model based on the em algorithm. *Biometrics 48*(3), 795–806.

Lachenbruch, P. A. (1998). Sensitivity, specificity, and vaccine efficacy. *Controlled clinical trials 19*(6), 569–574.

Lee, K. and D. S. Small (2018). Estimating the malaria attributable fever fraction accounting for parasites being killed by fever and measurement error. *Journal of the American Statistical Association* (just-accepted), 1–38.

Mabunda, S., J. J. Aponte, A. Tiago, and P. Alonso (2009). A country-wide malaria survey in mozambique. ii. malaria attributable proportion of fever and establishment of malaria case definition in children across different epidemiological settings. *Malaria Journal 8*(1), 74.

Mahmoudi, S. and H. Keshavarz (2017). Efficacy of phase 3 trial of rts, s/as01 malaria vaccine: the need for an alternative development plan. *Human vaccines & immunotherapeutics 13*(9), 2098–2101.

Marsaglia, G. et al. (2006). Ratios of normal variables. *Journal of Statistical Software 16*(4), 1–10.

Martens, E. P., W. R. Pestman, A. de Boer, S. V. Belitser, and O. H. Klungel (2006). Instrumental variables: application and limitations. *Epidemiology*, 260–267.

Martinussen, T., S. Vansteelandt, and P. K. Andersen (2018). Subtleties in the interpretation of hazard ratios. *arXiv preprint arXiv:1810.09192*.

Moorthy, V., Z. Reed, and P. G. Smith (2007). Measurement of malaria vaccine efficacy in phase iii trials: Report of a who consultation. *Vaccine 25*(28), 5115 – 5123.

Moorthy, V. S. and W. R. Ballou (2009). Immunological mechanisms underlying protection mediated by rts, s: a review of the available data. *Malaria journal 8*(1), 312.

Morton, N. (2001). Linkage disequilibrium. In S. Brenner and J. H. Miller (Eds.), *Encyclopedia of Genetics*, pp. 1105. New York: Academic Press.

Ndila, C. M., S. Uyoga, A. W. Macharia, G. Nyutu, N. Peshu, J. Ojal, M. Shebe, K. O. Awuondo, N. Mturi, B. Tsofa, et al. (2018). Human candidate gene polymorphisms and risk of severe malaria in children in kilifi, kenya: a case-control association study. *The Lancet Haematology 5*(8), e333–e345.

Neuwald, A. F. and P. Green (1994). Detecting patterns in protein sequences. *Journal of molecular biology 239*(5), 698–712.

North, T.-L., N. M. Davies, S. Harrison, A. R. Carter, G. Hemani, E. Sanderson, K. Tilling, and L. D. Howe (2019). Using genetic instruments to estimate interactions in mendelian randomization studies.

Olotu, A., G. Fegan, J. Wambua, G. Nyangweso, K. O. Awuondo, A. Leach, M. Lievens, D. Leboulleux, P. Njuguna, N. Peshu, and et al. (2013, Mar). Four-year efficacy of rts,s/as01e and its interaction with malaria exposure. *New England Journal of Medicine 368*(12), 1111–1120.

Patel, J. C., S. M. Taylor, S. R. Meshnick, K. L. Thwai, F. O. ter Kuile, C. Khairallah, L. Kalilani, and V. Mwapasa (2016, May). Absence of association between sickle trait

hemoglobin and placental malaria outcomes. *The American Journal of Tropical Medicine and Hygiene 94*(5), 1002–1007.

Penny, M. A., K. Galactionova, M. Tarantino, M. Tanner, and T. A. Smith (2015). The public health impact of malaria vaccine rts, s in malaria endemic africa: country-specific predictions using 18 month follow-up phase iii data and simulation models. *BMC medicine 13*(1), 170.

Penny, M. A., N. Maire, A. Studer, A. Schapira, and T. A. Smith (2008). What should vaccine developers ask? simulation of the effectiveness of malaria vaccines. *PLoS One 3*(9), e3193.

Rassen, J. A., M. A. Brookhart, R. J. Glynn, M. A. Mittleman, and S. Schneeweiss (2009). Instrumental variables i: instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships. *Journal of clinical epidemiology 62*(12), 1226–1232.

Rees, J. M. B., C. N. Foley, and S. Burgess (2019, 08). Factorial Mendelian randomization: using genetic variants to assess interactions. *International Journal of Epidemiology*. dyz161.

Regules, J. A., J. F. Cummings, and C. F. Ockenhouse (2011). The rts, s vaccine candidate for malaria. *Expert review of vaccines 10*(5), 589–599.

Rosenbaum, P. R. (2002). *Observational Studies.* New York: Springer.

Rosenbaum, P. R. (2011). Some approximate evidence factors in observational studies. *Journal of the American Statistical Association 106*(493), 285–295.

Rosenblum, M. and M. J. van der Laan (2009). Using regression models to analyze randomized trials: Asymptotically valid hypothesis tests despite incorrectly specified models. *Biometrics 65*(3), 937–945.

Rosenblum, M. and M. J. van der Laan (2010, Jan). Simple, efficient estimators of treatment effects in randomized trials using generalized linear models to leverage baseline variables. *The International Journal of Biostatistics 6*(1).

Rubin, D. B. (2005). Causal inference using potential outcomes. *Journal of the American Statistical Association 100*(469), 322–331.

Small, D. S., J. Cheng, and T. R. Ten Have (2010). Evaluating the efficacy of a malaria vaccine. *The international journal of biostatistics 6*(2).

Smith, G. D. and S. Ebrahim (2008). Mendelian randomization: genetic variants as instruments for strengthening causal inference in observational studies. In *Biosocial Surveys*. National Academies Press (US).

Smith, T., J. A. Schellenberg, and R. Hayes (1994). Attributable fraction estimates and case definitions for malaria in endemic. *Statistics in medicine 13*(22), 2345–2358.

Stensrud, M. J., M. Valberg, K. Røysland, and O. O. Aalen (2017, May). Exploring selection bias by causal frailty models. *Epidemiology 28*(3), 379–386.

Taylor, S. M., C. M. Parobek, and R. M. Fairhurst (2012). Haemoglobinopathies and the clinical epidemiology of malaria: a systematic review and meta-analysis. *The Lancet infectious diseases 12*(6), 457–468.

Ter Kuile, F. O., A. A. Lal, S. K. Kariuki, Y. P. Shi, L. B. Mirel, J. M. Vulule, P.-H. P. A., M. S. Kolczak, W. A. Hawley, B. L. Nahlen, and et al. (2003, Apr). Impact of permethrin-treated bed nets on malaria, anemia, and growth in infants in an area of intense perennial malaria transmission in western kenya. *The American Journal of Tropical Medicine and Hygiene 68*(4_suppl), 68–77.

The RTS,S Clinical Trials Partnership (2011, Nov). First results of phase 3 trial of rts,s/as01 malaria vaccine in african children. *New England Journal of Medicine 365*(20), 1863–1875.

The RTS,S Clinical Trials Partnership (2014, Jul). Efficacy and safety of the rts,s/as01 malaria vaccine during 18 months after vaccination: A phase 3 randomized, controlled trial in children and young infants at 11 african sites. *PLoS Medicine 11*(7), e1001685.

Therneau, T. M., P. M. Grambsch, and V. S. Pankratz (2003). Penalized survival models and frailty. *Journal of computational and graphical statistics 12*(1), 156–175.

Van der Laan, M. J. and S. Rose (2011). *Targeted learning: causal inference for observational and experimental data.* New York, NY: Springer Science & Business Media.

van der Laan, M. J. and D. Rubin (2006, Jan). Targeted maximum likelihood learning. *The International Journal of Biostatistics 2*(1).

VanderWeele, T. J. and M. A. Hernan (2013, Jan). Causal inference under multiple versions of treatment. *Journal of Causal Inference 1*(1), 1–20.

Wald, A. (1940, Sep). The fitting of straight lines if both variables are subject to error. *The Annals of Mathematical Statistics 11*(3), 284–300.

Wienke, A. (2010, Jul). Frailty models in survival analysis. *Chapman & Hall/CRC Biostatistics Series*.

Williams, T. N. (2011, Oct). How do hemoglobins s and c result in malaria protection? *Journal of Infectious Diseases 204*(11), 1651–1653.

Williams, T. N., T. W. Mwangi, S. Wambua, N. D. Alexander, M. Kortok, R. W. Snow, and K. Marsh (2005). Sickle cell trait and the risk of plasmodium falciparum malaria and other childhood diseases. *The Journal of infectious diseases 192*(1), 178–186.

World Health Organization (2018). *Who world malaria report 2018.* Luxembourg: World Health Organization.

Wu, Y., R. E. Sinden, T. S. Churcher, T. Tsuboi, and V. Yusibov (2015). Development of malaria transmission-blocking vaccines: from concept to product. In *Advances in parasitology*, Volume 89, pp. 109–152. Elsevier.