Evaluating the Long-Term Efficacy of COVID-19 Vaccines

Dan-Yu Lin,[1] Donglin Zeng,[1] and Peter B. Gilbert[2]

1. Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, USA;

2. Vaccine and Infectious Disease Division, Fred Hutch, Seattle, Washington, USA

Summary: We show how to estimate potentially waning long-term efficacy of COVID-19 vaccines in randomized, placebo-controlled clinical trials with staggered enrollment of participants and sequential crossover of placebo recipients.

Abstract: Large-scale deployment of safe and durably effective vaccines can halt the COVID-19 pandemic.[1−3] However, the high vaccine efficacy reported by ongoing phase 3 placebo-controlled clinical trials is based on a median follow-up time of only about two months[4−5] and thus does not pertain to long-term efficacy. To evaluate the duration of protection while allowing trial participants timely access to efficacious vaccine, investigators can sequentially cross placebo recipients to the vaccine arm according to priority groups. Here, we show how to estimate potentially time-varying placebo-controlled vaccine efficacy in this type of staggered vaccination of placebo recipients. In addition, we compare the performance of blinded and unblinded crossover designs in estimating long-term vaccine efficacy.

Keywords: booster vaccination; durability of vaccine efficacy; phase 3 trials; SARS-CoV-2; severe COVID-19; waning efficacy

Authors' Information: Dan-Yu Lin, Ph.D., is Dennis Gillings Distinguished Professor of Biostatistics, and Donglin Zeng, Ph.D., is Professor of Biostatistics, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC 27599-7420, USA. Peter B. Gilbert, Ph.D., is Member, Vaccine and Infectious Disease Division, Fred Hutch, Seattle, WA 98109-1024, USA.

Correspondence: Dan-Yu Lin, PhD, Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599-7420, USA (lin@bios.unc.edu).

Introduction

A number of studies have been conducted around the world to evaluate the efficacy and safety of investigational vaccines against novel coronavirus disease-2019 (COVID-19).[4–12] Interim results from several large-scale phase 3 randomized, observer-blinded, placebo-controlled clinical trials have demonstrated high vaccine efficacy,[4–6] far exceeding the FDA and WHO thresholds of 50% reduction of symptomatic disease.[13–14] However, those trials have been underway for only several months, and the data they have collected thus far can only speak to short-term vaccine efficacy. For example, the interim results for the Pfizer/BioNTech vaccine BNT162b2 and the Moderna vaccine mRNA-1273 that were presented for regulatory approval are based on a median follow-up time of approximately two months after the second dose;[4–5] therefore, the reported 94-95% efficacy pertains only to an average of two months post vaccination.

Vaccine efficacy can wane over time because of declining immunologic memory or changing antigenicity of the pathogen. Thus, it is critically important to determine the duration of protection and the need for additional booster doses of vaccine. Certainly, after FDA issues Emergency Use Authorization (EUA), vaccine sponsors should continue to collect placebo-controlled data in any ongoing trials for as long as feasible.[15]

Although continuing blinded follow-up of the original treatment arms offers the ideal opportunity to evaluate long-term vaccine efficacy and safety, placebo recipients should be offered the vaccine at some point after an EUA. One strategy is the "rolling crossover", which vaccinates placebo recipients around the same time as general population members in the same priority tier. Under this design, placebo participants are vaccinated at different times, with the timing of vaccination depending on enrollment characteristics that define their priority tier.

In this article, we show how to evaluate long-term vaccine efficacy under the rolling crossover design. Because of staggered enrollment, existing statistical methods do not produce valid estimates of time-varying vaccine efficacy even if the placebo group is maintained

throughout the study. Our approach properly accounts for staggered enrollment and change of background incidence over time, as well as for the fact that higher-risk placebo volunteers may get vaccinated earlier than lower-risk ones during the crossover period. It provides unbiased estimation of the entire curve of placebo-controlled vaccine efficacy as a function of time elapsed since vaccination, up to the point where most of the placebo recipients have crossed over to the vaccine arm. We investigate the bias and precision in estimating long-term vaccine efficacy under various blinded and unblinded crossover designs and discuss how to perform sensitivity analysis when unblinded follow-up data are used.

Methods

Figure 1 provides a schematic illustration of the rolling crossover strategy in the context of the two mRNA vaccine trials. In this scenario, participants were screened and randomly assigned to vaccine or placebo over a 4-month period, and the vaccine was approved in the 5th month on the basis of interim results. Under the proposed design, crossover based on priority groups will occur over the next five months.

The endpoint of interest is time to symptomatic COVID-19 disease. We allow the background risk of disease to vary over the calendar time and to depend on baseline risk factors, such as age, gender, ethnicity, race, occupation, and underlying health conditions. In addition, we allow the relative risk of disease for vaccine versus placebo to depend on the duration elapsed since vaccination. We define the τ-day vaccine efficacy as the percentage reduction in the cumulative incidence of disease over the τ-day period for those who were vaccinated at the start of the τ-day period versus those who have not been vaccinated by day τ.

In Supplemental Appendix 1, we formulate the above framework through an adaptation of the popular Cox proportional hazards model[16−17] and derive the maximum likelihood estimator for the curve of placebo-controlled vaccine efficacy as a function of time elapsed since vaccination. In addition, we show that the estimator is approximately unbiased and normally distributed, with a variance that can be estimated analytically, making it possible to construct valid confidence intervals for the vaccine-efficacy curve.

Results

We conducted a series of simulation studies mimicking the BNT162b2 vaccine trial. We considered 40,000 participants, who entered the trial at a constant rate over a 4-month period and were randomly assigned to vaccine or placebo in a 1:1 ratio. The vaccine received an EUA from FDA at the 5th month, by which time there were about 300 COVID-19 cases in the placebo group. To reflect the increase of COVID-19 cases since last summer and the expected downward trend in the spring due to herd immunity, we let the disease risk increase over the first 7 months and decrease afterward. We set the vaccine efficacy at 5 months to be 95% and the vaccine efficacy at 10 months to be 95%, 50%, or 0%.

We considered the statistically optimal design of keeping all participants on their original treatment assignments until the end of the trial, which was set to be slightly over 10 months. We refer to this design as Plan A, which serves as a benchmark. We also considered three blinded crossover designs:

B. Crossover occurs over months 6-10 according to priority tier groups.

C. 20% of participants follow Plan A, and the rest follow Plan B.

D. Crossover occurs for all participants in month 6.

Both B and C are priority-tier dependent rolling crossover designs. The difference is that under Plan B, all placebo recipients cross over to the vaccine arm, whereas under Plan C, 20% of participants choose for altruistic reasons to stay on their original treatment assignments. For Plan D, all placebo recipients are vaccinated over the course of a month without any priority tiering. With blinded crossover, placebo participants receive the vaccine and vaccine participants receive the placebo at the point of crossover, and all participants are followed until the end of the trial, which is 10.5 months since trial initiation. The designs of these simulation studies are detailed in Supplemental Appendix 2.

The results, based on 10,000 simulated datasets, are summarized in Table 1 and Figure 2. The proposed method yields virtually unbiased estimates of the vaccine-efficacy curves over the 10-month period for Plans A–C in all three scenarios (95%, 50%, 0%) of long-term

vaccine efficacy; it also yields accurate variance estimates, such that the confidence intervals have proper coverage probabilities. When vaccine efficacy is constant over time, the standard errors for the estimates of vaccine efficacy under Plans B and C are slightly lower than those of Plan A. When vaccine efficacy wanes over time, the standard errors for the estimates of 5-month vaccine efficacy under Plans B and C are also slightly lower than those of Plan A; however, the standard errors for the estimates of 10-month vaccine efficacy under Plans B and C are higher than those of Plan A, with the standard errors being slightly lower under Plan C than under Plan B. Under Plan D, the estimates of 10-month vaccine efficacy are slightly biased, with high standard errors; these results are not surprising, because under this plan, the number of unvaccinated participants diminishes rapidly after month 6.

We also evaluated the performance of the standard Cox regression,[16–17] with vaccine status as a potentially time-dependent covariate. Because it estimates an overall vaccine efficacy among individuals who have been vaccinated for different amounts of time, this method over-estimates long-term vaccine efficacy when vaccine efficacy decreases over time, although the estimation is unbiased when vaccine efficacy is constant over time.

We conducted a second series of simulation studies by considering unblinded crossover, under which participants are notified of their original treatment assignments and placebo recipients are vaccinated soon after. In this series, crossover occurs at the same time as in Plans B–D but is unblinded; we refer to these three plans as B'–D' (Supplementary Appendix 2). Because vaccine recipients may engage in riskier behavior upon unblinding, we discarded the data collected after unblinding for both the vaccine and placebo groups by censoring each participant's time to disease at their time of unblinding visit; the resulting analysis avoids any bias due to behavioral confounding.

The results for the second series of simulation studies, based on 10,000 simulated datasets, are summarized in Table 2 and Figure 3. The proposed method yields unbiased estimates of the vaccine-efficacy curves under Plans B' and C', but with much higher standard errors for estimates of long-term vaccine efficacy than under Plans B and C; the standard errors can be accurately estimated, although only up to month 10. Under Plan D', the estimates of

5-month vaccine efficacy are unbiased; however, the estimates of 10-month vaccine efficacy are biased, especially when vaccine efficacy wanes over time. Naturally, long-term efficacy cannot be estimated as well under unblinded crossover as under blinded crossover because of limited blinded follow-up data, especially under Plan D'. Of note, the bias of the standard Cox regression in estimating waning vaccine efficacy is more severe under unblinded crossover than under blinded crossover.

Figure 4 displays the analysis results produced by the proposed methods for one of the trials simulated with 50% 10-month vaccine efficacy. The estimates of vaccine efficacy are close to the truth, and the 95% confidence intervals cover the truth mostly — up to the end of blinded follow-up. In terms of estimating long-term vaccine efficacy, Plan C is nearly as good as Plan A and is slightly better than Plan B, which is better than Plan D; Plans B, C, and D are substantially better than Plans B', C', and D', respectively.

Discussion

For a preventive COVID-19 vaccine to be administered to millions of people, including healthy individuals, its safety and efficacy must be demonstrated in a clear and compelling manner. Although preliminary results from ongoing phase 3 clinical trials have revealed higher than expected efficacy of COVID-19 vaccines,[4–6] additional follow-up is required to assess long-term efficacy and safety. Indeed, FDA does not consider issuance of an EUA, in and of itself, as grounds for stopping blinded follow-up in an ongoing clinical trial.[15]

We recommend the rolling crossover design, which allows placebo recipients to be vaccinated in a timely manner while still making it possible to assess long-term safety and efficacy. The standard Cox regression seriously over-estimates long-term vaccine efficacy when vaccine efficacy wanes over time, as shown in our simulation studies. We have developed a valid and efficient approach to evaluate the potential waning efficacy of a COVID-19 vaccine. The estimated curve of time-varying vaccine efficacy can be used to determine when a booster vaccination is needed to sustain protection; this information is also an important input parameter in mathematical modeling of the population impact of COVID-19 vaccines.

To ensure high-quality follow-up data, crossover should ideally be blinded, with placebo participants receiving vaccine and vaccine participants receiving placebo at the time of crossover. Unblinded crossover has practical benefits over blinded crossover: it reduces operational complexity and trial cost. However, unblinding can lead to differential exposure to SARS-CoV-2 between the original vaccinees and the placebo crossovers, which in turn can bias the estimation of vaccine efficacy.

This bias can be avoided by analyzing only the blinded follow-up data. However, discarding the unblinded follow-up data may substantially reduce the precision in estimating long-term vaccine efficacy. We suggest to estimate vaccine efficacy in two ways, one using all follow-up data and one using only blinded follow-up data, and compare the two results.

Another solution is to apply our methods to all follow-up data but perform a sensitivity analysis to assess the robustness of the results to potential unmeasured confounding caused by unblinding of trial participants. In Supplemental Appendix 3, we show how to apply a best-practice general methodology in epidemiological research[18−19] to perform this sensitivity analysis. Specifically, we can assess how strong unmeasured confounding due to unblinding would need to be in order to fully explain away the observed vaccine efficacy. We can also provide a conservative estimate of vaccine efficacy that accounts for unmeasured confounding.

Recently, Follmann et al [20] advocated blinded crossover and continued follow-up of trial participants to assess vaccine durability and potential delayed enhancement of disease. They estimated an overall vaccine efficacy for the original vaccine arm in the post-crossover period (after all placebo recipients have been vaccinated) by imputing the case count for a counterfactual placebo group under several assumptions: that the vaccine and placebo groups remain balanced at the point of cross-over; that the crossover pattern and duration of follow-up in the second period are comparable to the enrollment pattern and duration of follow-up in the first period; and that the benefit of vaccination over time has the same profile between the original vaccinees and the placebo crossovers. We address the issue of long-term vaccine efficacy in a different way, using observed data to estimate the entire curve of vaccine efficacy as a function of time elapsed since vaccination, up to the point where most participants have

been vaccinated. Our approach requires minimal assumptions and is applicable to both blinded and unblinded crossover plans, with any length of additional follow-up.

Acknowledgments

The authors thank Yu Gu and Bridget I. Lin for assistance. This work was supported by the National Institutes of Health grants R01 AI029168, R01 GM124104, P01 CA142538, and UM1 AI068635.

References

1. Krause P, Fleming, TR, Longini I, et al (2020). COVID-19 vaccine trials should seek worthwhile efficacy. Lancet, 396: 741-743.

2. Cohen MS, Corey L (2020). Combination prevention for COVID-19. Science 368: 551.

3. Corey L, Mascola JR, Fauci AS, Collins FS (2020). A strategic approach to COVID-19 vaccine R&D. Science 368: 948-50.

4. Polack FP, Thomas SJ, Kitchin N, et al (2021). Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. New Eng J Med, 383: 2603-15.

5. Baden LR, El Sahly HM, Essink B, et al (2021). Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. New Eng J Med (DOI: 10.1056/NEJMoa2035389).

6. Voysey M, Clemens SAC, Madhi SA, et al (2021). Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. Lancet, in press.

7. Xia S, Duan K, Zhang Y, et al (2020). Effect of an inactivated vaccine against SARS-CoV-2 on safety and immunogenicity outcomes: Interim analysis of 2 randomized clinical trials. J Am Med Ass, 324: 951-960.

8. Jackson LA, Anderson EJ, Rouphael NG, et al (2020). An mRNA vaccine against SARS-CoV-2 – Preliminary report. New Eng J Med, 383: 1920-31.

9. Folegatti PM, Ewer KJ, Aley PK, et al (2020). Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. Lancet, 396: 467-478.

10. Keech C, Albert G, Cho I, et al (2020). Phase 1-2 Trial of a SARS-CoV-2 recombinant spike protein nanoparticle vaccine. New Eng J Med, 383: 2320-32.

11. The World Health Organization (2020). Draft landscape of COVID-19 candidate vaccines. https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines (accessed December 8, 2020).

12. The World Health Organization (2020). An international randomised trial of candidate vaccines against COVID-19. https://www.who.int/publications-detail/an-international-randomised-trial-of-candidate-vaccines-against-covid-19.

13. U.S. Food and Drug Administration (2020). Development and Licensure of Vaccines to Prevent COVID-19: Guidance for Industry.

14. The World Health Organization (2020). WHO target product profiles for COVID-19 vaccines. https://www.who.int/who-documents-detail/who-target-productprofiles-for-covid-19-vaccines.

15. U.S. Food and Drug Administration (2020). Emergency Use Authorization for Vaccines to Prevent COVID-19: Guidance for Industry. https://www.fda.gov/media/142749/download.

16. Cox DR (1972). Regression models and life-tables. J Roy Stat Soc B, 34: 187-202.

17. Kalbfleisch JD, Prentice RL (2011). The Statistical Analysis of Failure Time Data, 2nd Ed. John Wiley & Sons.

18. VanderWeele TJ, Ding P (2017). Sensitivity analysis in observational research: introducing the E-value. Ann Int Med, 167: 268-74.

19. Ding P, VanderWeele TJ (2016). Sensitivity analysis without assumptions. Epidemiology, 27: 368.

20. Follmann D, Fintzi J, Fay MP, et al. (2020). Assessing durability of vaccine effect following blinded crossover in COVID-19 vaccine efficacy trials. https://www.medrxiv.org/content/10.1101/2020.12.14.20248137v1

Table 1. Estimation of Vaccine Efficacy (VE) by the Proposed and Standard Methods Under No Crossover (A) and Three Blinded Crossover Plans (B–D) in Three Scenarios of Long-Term Vaccine Efficacy

			True	Proposed method				Standard	
Scenario	Plan	Time	VE	Mean	SE	SEE	CP	Mean	SE
(a) 5-mos VE = 95%	A	5 mos	95%	95.0%	0.8%	0.8%	95.2%		
10-mos VE = 95%		10 mos	95%	95.0%	0.8%	0.8%	94.5%	95.0%	0.6%
	B	5 mos	95%	95.0%	0.7%	0.7%	95.1%		
		10 mos	95%	95.1%	0.8%	0.7%	94.5%	95.0%	0.6%
	C	5 mos	95%	95.0%	0.7%	0.7%	95.2%		
		10 mos	95%	95.0%	0.7%	0.7%	94.5%	95.0%	0.6%
	D	5 mos	95%	95.0%	0.8%	0.8%	94.8%		
		10 mos	95%	95.2%	0.9%	0.9%	93.6%	95.0%	0.7%
(b) 5-mos VE = 95%	A	5 mos	95%	95.0%	0.8%	0.8%	95.3%		
10-mos VE = 50%		10 mos	50%	49.9%	4.7%	4.6%	95.0%	73.5%	1.6%
	B	5 mos	95%	95.0%	0.7%	0.7%	95.0%		
		10 mos	50%	49.9%	6.4%	6.4%	95.3%	84.2%	1.0%
	C	5 mos	95%	95.0%	0.7%	0.7%	95.2%		
		10 mos	50%	49.9%	5.7%	5.6%	95.2%	81.2%	1.1%
	D	5 mos	95%	95.0%	0.8%	0.8%	94.7%		
		10 mos	50%	49.3%	10.0%	10.0%	95.1%	92.8%	0.8%
(c) 5-mos VE = 95%	A	5 mos	95%	95.0%	0.8%	0.8%	95.2%		
10-mos VE = 0%		10 mos	0%	-0.2%	8.1%	8.0%	95.1%	54.3%	2.3%
	B	5 mos	95%	95.0%	0.7%	0.7%	95.2%		
		10 mos	0%	-0.3%	11.5%	11.6%	94.8%	77.0%	1.2%
	C	5 mos	95%	95.0%	0.7%	0.7%	95.3%		
		10 mos	0%	-0.4%	10.3%	10.2%	94.6%	71.0%	1.4%
	D	5 mos	95%	95.0%	0.8%	0.8%	94.6%		
		10 mos	0%	-1.4%	18.8%	18.7%	95.2%	91.5%	0.8%

Note: Mean and SE denote the mean and standard error of the estimator for vaccine efficacy, SEE denotes the mean of the standard error estimator, and CP denotes the coverage probability of the 95% confidence interval.

Table 2. Estimation of Vaccine Efficacy (VE) by the Proposed and Standard Methods Under Three Unblinded Crossover Plans (B'–D') in Three Scenarios of Long-Term Vaccine Efficacy

			True	Proposed method				Standard	
Scenario	Plan	Time	VE	Mean	SE	SEE	CP	Mean	SE
(a) 5-mos VE = 95%	B'	5 mos	95%	95.0%	0.9%	0.9%	95.3%		
		8 mos	95%	95.0%	0.9%	0.9%	94.0%		
10-mos VE = 95%		10 mos	95%	95.1%	2.0%	1.3%	85.0%	95.0%	0.8%
	C'	5 mos	95%	95.0%	0.9%	0.9%	95.1%		
		8 mos	95%	95.0%	0.9%	0.9%	94.2%		
		10 mos	95%	94.9%	7.1%	1.5%	83.9%	95.0%	0.8%
	D'	5 mos	95%	95.0%	1.0%	1.0%	94.6%		
		8 mos	95%	96.3%	1.3%	0.9%	65.1%		
		10 mos	95%	97.0%	1.1%	0.7%	29.7%	95.0%	1.0%
(b) 5-mos VE = 95%	B'	5 mos	95%	95.0%	0.9%	0.9%	95.1%		
		8 mos	81%	80.8%	2.8%	2.8%	95.1%		
10-mos VE = 50%		10 mos	50%	52.1%	21.4%	14.3%	86.4%	88.2%	1.2%
	C'	5 mos	95%	95.0%	0.9%	0.9%	95.3%		
		8 mos	81%	80.8%	2.8%	2.8%	94.9%		
		10 mos	50%	50.1%	20.6%	14.7%	88.1%	88.1%	1.3%
	D'	5 mos	95%	95.0%	1.2%	1.1%	94.8%		
		8 mos	81%	94.2%	2.3%	1.7%	5.2%		
		10 mos	50%	95.4%	1.9%	1.4%	0.4%	95.5%	0.9%
(c) 5-mos VE = 95%	B'	5 mos	95%	95.0%	0.9%	0.9%	95.1%		
		8 mos	71%	71.1%	3.8%	3.9%	95.2%		
10-mos VE = 0%		10 mos	0%	3.3%	34.7%	26.0%	88.5%	83.4%	1.4%
	C'	5 mos	95%	95.0%	0.9%	0.9%	95.2%		
		8 mos	71%	71.0%	3.8%	3.7%	94.8%		
		10 mos	0%	-0.5%	35.2%	26.1%	89.8%	83.0%	1.5%
	D'	5 mos	95%	95.0%	1.2%	1.2%	94.9%		
		8 mos	71%	93.3%	3.0%	2.1%	3.3%		
		10 mos	0%	94.7%	2.4%	1.7%	0.2%	95.5%	0.9%

Note: Mean and SE denote the mean and standard error of the estimator for vaccine efficacy, SEE denotes the mean of the standard error estimator, and CP denotes the coverage probability of the 95% confidence interval.

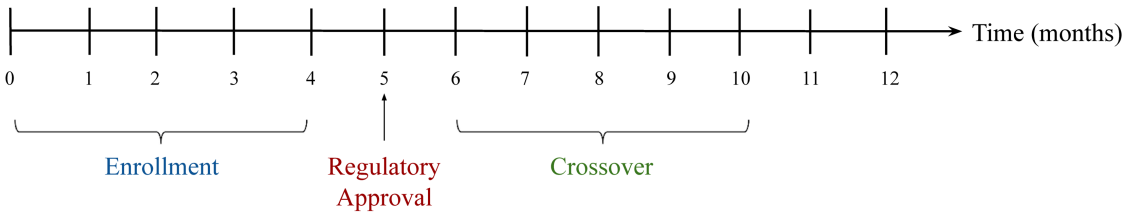

Figure 1. Possible timeline for a phase 3 COVID-19 vaccine trial.

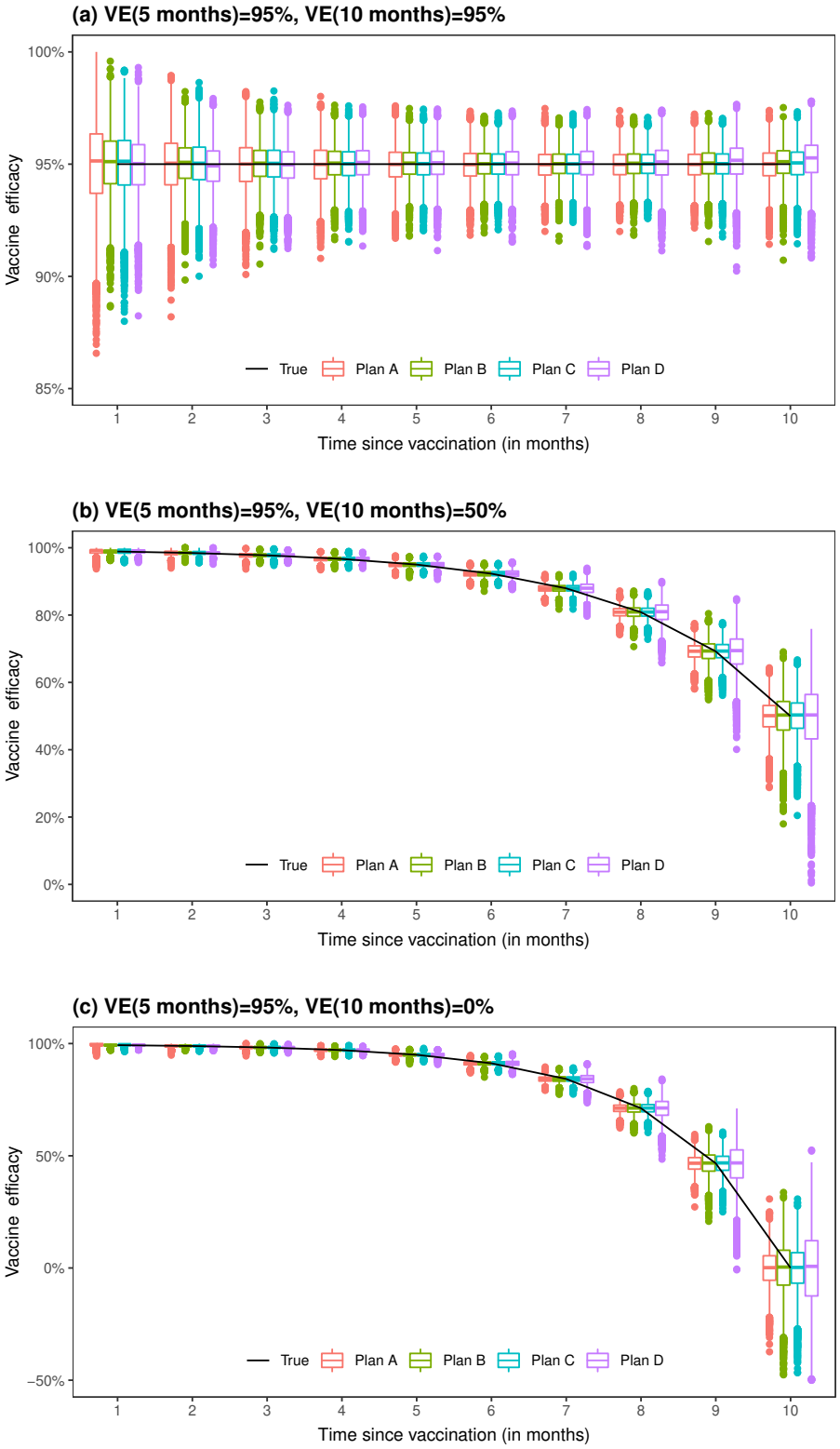

Figure 2. Proposed estimates of vaccine efficacy under no crossover and three blinded crossover plans in three scenarios of long-term vaccine efficacy.

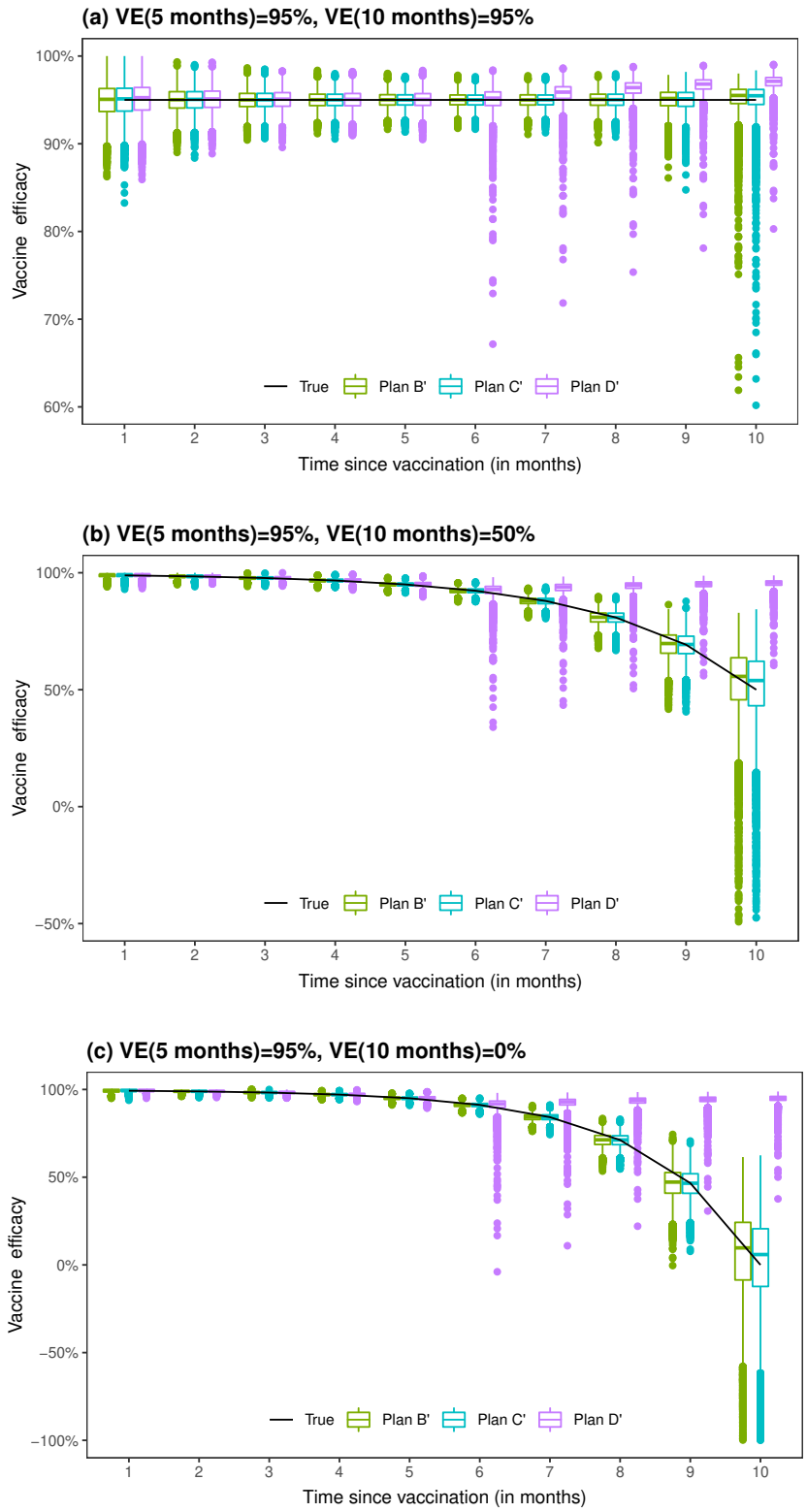

Figure 3. Proposed estimates of vaccine efficacy under three unblinded crossover plans in three scenarios of long-term vaccine efficacy.

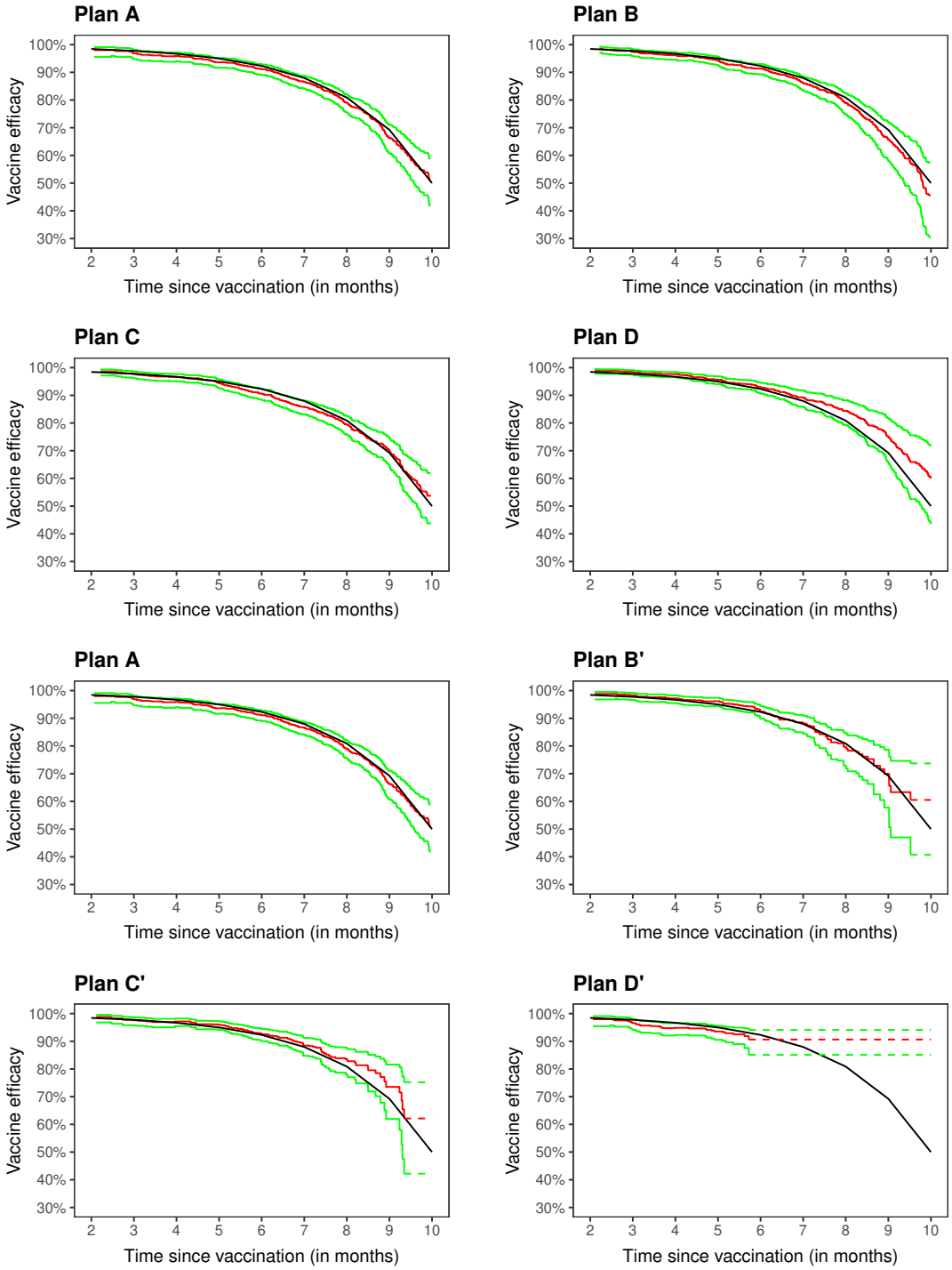

Figure 4. Estimation of vaccine efficacy in a clinical trial with no crossover (A), three blinded crossover (B–D), and three unblinded crossover (B'–D'): the black curve pertains to the true value, the red curve to the proposed estimate, and the green curves to the 95% confidence intervals.

Supplemental Appendix

1. Estimation of Vaccine Efficacy

Let T and S denote the time to symptomatic COVID-19 and the time to vaccination, respectively; both times are measured in days from the start of the clinical trial. In addition, let X denote baseline risk factors (e.g., age, occupation, comorbidities). We specify that the hazard function of T is related to S and X through a Cox proportional hazards model with a time-dependent regression coefficient

$$\lambda(t|S, X) = \lambda_0(t)v(t - S)^{I(t>S)}e^{\beta^{\mathrm{T}}X}, \tag{1}$$

where $\lambda_0(\cdot)$ is an arbitrary baseline hazard function, $v(\cdot)$ is a positive function characterizing the time-varying effect of vaccination, β is a set of regression parameters representing the effects of baseline risk factors, and $I(\cdot)$ is the indicator function. Under this formulation, the effect of vaccine on the risk of disease depends on the time elapsed since vaccination, but not on the specific date of vaccination or on the baseline risk factors.

One may define the t-day vaccine efficacy (VE) as the proportionate reduction in the cumulative incidence of disease at time t for individuals who have been vaccinated for t days:

$$VE(t) = 1 - \frac{\Pr(T \leq t|S = 0, X)}{\Pr(T \leq t|S > t, X)} = 1 - \frac{1 - \exp\{-e^{\beta^{\mathrm{T}}X}\int_0^t v(u)\lambda_0(u)du\}}{1 - \exp\{-e^{\beta^{\mathrm{T}}X}\int_0^t \lambda_0(u)du\}},$$

which is approximately

$$1 - \frac{\int_0^t v(u)\lambda_0(u)du}{\int_0^t \lambda_0(u)du}$$

for rare disease. If $\lambda_0(\cdot)$ is approximately constant, then

$$VE(t) = 1 - V(t)/t, \tag{2}$$

where

$$V(t) = \int_0^t v(u)du.$$

For simplicity, we define the time-varying vaccine efficacy by equation (2).

Suppose that the clinical trial enrolls a total of n participants. For $i = 1, \ldots, n$, let R_i, T_i, S_i, C_i, and X_i denote the entry time, the time to symptomatic COVID-19, the time to vaccination, the time to loss to follow-up, and the baseline risk factors for the ith participant. The data consist of $(R_i, Y_i, \Delta_i, D_i, D_i S_i, X_i)$ $(i = 1, \ldots, n)$, where $Y_i = \min(T_i, C_i)$, $\Delta_i = I(T_i \leq C_i)$, and $D_i = I(S_i \leq Y_i)$.

We assume that (R_i, S_i, C_i) are independent of T_i conditional on X_i. The likelihood takes the form

$$\prod_{i=1}^{n} \left\{ \lambda_0(Y_i) e^{\beta^{\mathrm{T}} X_i} \right\}^{\Delta_i} v(\widetilde{Y}_i)^{D_i \Delta_i}$$

$$\times \exp\left(-e^{\beta^{\mathrm{T}} X_i} \left[(1 - D_i) \int_{R_i}^{Y_i} \lambda_0(t) dt + D_i \left\{ \int_{R_i}^{S_i} \lambda_0(t) dt + \int_{S_i}^{Y_i} \lambda_0(t) dV(t - S_i) \right\} \right] \right)$$

$$= \prod_{i=1}^{n} \left\{ \lambda_0(Y_i) e^{\beta^{\mathrm{T}} X_i} \right\}^{(1-D_i)\Delta_i} \exp\left\{ -e^{\beta^{\mathrm{T}} X_i} \int_{R_i}^{Y_i^*} \lambda_0(t) dt \right\}$$

$$\times \left[\left\{ \lambda_0(Y_i) e^{\beta^{\mathrm{T}} X_i} v(\widetilde{Y}_i) \right\}^{\Delta_i} \exp\left\{ -e^{\beta^{\mathrm{T}} X_i} \int_{S_i}^{Y_i} \lambda_0(t) dV(t - S_i) \right\} \right]^{D_i},$$

where $\widetilde{Y}_i = Y_i - S_i$, and $Y_i^* = (1 - D_i) Y_i + D_i S_i = \min(Y_i, S_i)$. We approximate $\log \lambda_0(t)$ through splines with m basis functions, $B_1(t), \ldots, B_m(t)$, such that $\log \lambda_0(t) \approx \sum_{k=1}^{m} \gamma_k B_k(t)$. Let $\theta = (\beta, \gamma_1, \ldots, \gamma_m)^{\mathrm{T}}$, and $Z_i(t) = [X_i, B_1(t), \ldots, B_m(t)]^{\mathrm{T}}$. We perform the nonparametric maximum likelihood estimation,[1] in which $V(\cdot)$ is treated as a step function jumping at the time points $\widetilde{Y}_i$ with $D_i = \Delta_i = 1$. Thus, we maximize the objective function

$$\prod_{i=1}^{n} e^{(1-D_i)\Delta_i \theta^{\mathrm{T}} Z_i(Y_i)} \exp\left\{ -\int_{R_i}^{Y_i^*} e^{\theta^{\mathrm{T}} Z_i(t)} dt \right\} \left[\left(e^{\theta^{\mathrm{T}} Z_i(Y_i)} V\{\widetilde{Y}_i\} \right)^{\Delta_i} \exp\left\{ -\int_0^{\widetilde{Y}_i} e^{\theta^{\mathrm{T}} Z_i(t+S_i)} dV(t) \right\} \right]^{D_i}, \tag{3}$$

where $V\{t\}$ is the jump size of $V(\cdot)$ at t.

We first maximize the objective function in (3) for fixed θ to yield

$$V(t) = \sum_{i=1}^{n} \frac{D_i \Delta_i I(\widetilde{Y}_i \leq t)}{S^{(0)}(\theta; \widetilde{Y}_i)}. \tag{4}$$

Here and in the sequel, $S^{(k)}(\theta; y) = \sum_{j=1}^{n} D_j I(\widetilde{Y}_j \geq y) e^{\theta^{\mathrm{T}} Z_j(y+S_j)} Z_j(y+S_j)^{\otimes k}$, where $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$, and $a^{\otimes 2} = a a^{\mathrm{T}}$ for a column vector a. After plugging (4) into (3), we obtain the

profile likelihood for θ. Differentiating the profile log-likelihood with respect to θ yields the estimating function

$$\sum_{i=1}^{n}\left\{(1-D_i)\Delta_i Z_i(Y_i)-\int_{R_i}^{Y_i^*} e^{\theta^{\mathrm{T}} Z_i(t)} Z_i(t)dt\right\}+\sum_{i=1}^{n} D_i\Delta_i\left\{Z_i(Y_i)-\frac{S^{(1)}(\theta;\widetilde{Y}_i)}{S^{(0)}(\theta;\widetilde{Y}_i)}\right\}.$$

Denote the resulting estimator of θ by $\widehat{\theta}$. Replacing θ in (4) by $\widehat{\theta}$ yields the estimator of $V(t)$

$$\widehat{V}(t)=\sum_{i=1}^{n}\frac{D_i\Delta_i I(\widetilde{Y}_i\le t)}{S^{(0)}(\widehat{\theta};\widetilde{Y}_i)},$$

which is reminiscent of the Breslow estimator[2] for the cumulative baseline hazard function under the standard Cox model. We then estimate $VE(t)$ through equation (2).

Using counting-process martingale theory and other mathematical arguments,[3–4] we can show that $\widehat{V}(t)$ is consistent and asymptotically normal. In addition, the variance of $\widehat{V}(t)$ can be consistently estimated by $n^{-1}\sum_{i=1}^{n} W_i(t)^2$, where

$$W_i(t)=D_i\left[\frac{\Delta_i I(\widetilde{Y}_i\le t)}{S^{(0)}(\widehat{\theta};\widetilde{Y}_i)}-\sum_{j=1}^{n}\frac{D_j\Delta_j I\{\widetilde{Y}_j\le \min(t,\widetilde{Y}_i)\}e^{\widehat{\theta}^{\mathrm{T}} Z_i(\widetilde{Y}_j+S_i)}}{S^{(0)}(\widehat{\theta};\widetilde{Y}_j)^2}\right]$$
$$-\sum_{j=1}^{n}\frac{D_j\Delta_j I(\widetilde{Y}_j\le t)S^{(1)}(\widehat{\theta};\widetilde{Y}_j)^{\mathrm{T}}}{S^{(0)}(\widehat{\theta};\widetilde{Y}_j)^2}\,Q_i,$$

and

$$Q_i=\left[\sum_{i=1}^{n}\int_{R_i}^{Y_i^*} e^{\widehat{\theta}^{\mathrm{T}} Z_i(t)} Z_i(t)^{\otimes 2}dt+\sum_{i=1}^{n} D_i\Delta_i\left\{\frac{S^{(2)}(\widehat{\theta};\widetilde{Y}_i)}{S^{(0)}(\widehat{\theta};\widetilde{Y}_i)}-\frac{S^{(1)}(\widehat{\theta};\widetilde{Y}_i)^{\otimes 2}}{S^{(0)}(\widehat{\theta};\widetilde{Y}_i)^2}\right\}\right]^{-1}$$
$$\times\left[(1-D_i)\Delta_i Z_i(Y_i)-\int_{R_i}^{Y_i^*} e^{\widehat{\theta}^{\mathrm{T}} Z_i(t)} Z_i(t)dt+D_i\Delta_i\left\{Z_i(Y_i)-\frac{S^{(1)}(\widehat{\theta};\widetilde{Y}_i)}{S^{(0)}(\widehat{\theta};\widetilde{Y}_i)}\right\}\right.$$
$$\left.-\sum_{j=1}^{n}\frac{D_j\Delta_j I(\widetilde{Y}_i\ge\widetilde{Y}_j)e^{\widehat{\theta}^{\mathrm{T}} Z_i(\widetilde{Y}_j+S_i)}}{S^{(0)}(\widehat{\theta};\widetilde{Y}_j)}\left\{Z_i(\widetilde{Y}_j+S_i)-\frac{S^{(1)}(\widehat{\theta};\widetilde{Y}_j)}{S^{(0)}(\widehat{\theta};\widetilde{Y}_j)}\right\}\right].$$

These results can be used to construct confidence intervals for $VE(t)$. Specifically, we first construct the confidence intervals for $\log\{V(t)\}$ and then transform them to $VE(t)$. We can also show the convergence of $\widehat{V}(\cdot)$ to a Gaussian process and construct simultaneous confidence bands for $VE(\cdot)$.[4]

We have defined vaccine efficacy in terms of reduction in the cumulative incidence of disease. An alternative definition is reduction in the risk of disease.[5] Specifically, we define

the vaccine efficacy at time t by the proportionate reduction in the hazard rate at time t, i.e., $ve(t) = 1 - v(t)$. We refer to $VE(t)$ as t-day vaccine efficacy and to $ve(t)$ as day-t vaccine efficacy, the former pertaining to cumulative vaccine effect and the latter to instantaneous vaccine effect. If the hazard ratio is constant over time, then the two definitions are equivalent. If the hazard ratio increases over time, then $VE(t)$ is higher than $ve(t)$. We can estimate $v(t)$ and thus $ve(t)$ through kernel smoothing of $\widehat{V}(t)$.[6]

2. Designs of Simulation Studies

We assumed that 40,000 participants entered the study at a constant rate over four months, i.e., $R \sim \text{Uniform}(0, 4)$. We created a composite baseline risk score X, which takes values 1, 2, 3, 4, and 5 with equal probability. At study entry, half of the participants were assigned to vaccine and half to placebo. The statistically optimal design would be to maintain the original vaccine and placebo groups until the end of the study, which was set to be 10.5 months. We refer to this design as Plan A. We also considered three blinded crossover designs, under which placebo participants receive vaccine and vaccine participants receive placebo at the time of crossover and all participants are followed until the end of the study:

Plan B. Crossover occurs at month $(11 - X + G)$, where G follows the exponential distribution with mean of 0.5.

Plan C. 20% of participants follow Plan A and 80% follow Plan B.

Plan D. Crossover occurs at month $6 + G$, where G follows the exponential distribution with mean of 0.5.

Plan C mimics a scenario where all participants are offered the option of crossover but a small percentage (20%) choose to stay on the original assignments for the sake of science.

In addition, we considered three unblinded crossover designs:

Plan B'. Crossover occurs at month $(11.5 - X)$.

Plan C'. 20% of participants follow Plan A and 80% follow Plan B'.

Plan D'. Crossover occurs at month 6.5.

Under unblinded crossover, participants are notified of their original assignments at the time of crossover, and placebo participants receive the vaccine soon after. In practice, only placebo participants would cross over, since there is no need to give placebo to vaccine recipients after unblinding. In Plans B'–D', the time of crossover is the time of unblinding rather than the time when placebo participants actually receive the vaccine. Because participants might change their behavior upon discovering their original treatment assignments, we discarded the follow-up data collected after unblinding by censoring each participant's time to disease at their time of unblinding. This strategy avoids bias due to behavioral confounding, at the cost of reduced statistical efficiency.

We generated the event time T from model (1) with $\gamma = 0.2$,

$$\log \lambda_0(t) = -5.93 + 0.1t - 0.3(t-7)_+,$$

and

$$\log v(t) = a + bt, \quad t > 0,$$

where a and b were chosen to achieve the desired VE values at 5 and 10 months. We censored T at the time of unblinding crossover under Plans B'–D'.

For each simulated dataset, we estimated $\log \lambda_0(t)$ using a piece-wise constant function with 20 pieces placed at the equal quantiles of the observed event times. We then estimated $V(t)$ using the proposed method and estimated $VE(t)$ according to equation (2). For comparisons, we also fit a Cox model that includes X and time-dependent covariate $I(S < t)$ to estimate VE by 1 minus the estimated hazard ratio of the time-dependent covariate.

3. Sensitivity Analysis

We suggest reporting the E-value[7] as a summary measure of the evidence against the null hypothesis $H_0 : VE(t) \leq 0$. The E-value is the minimum strength of association on the risk ratio scale, i.e, $RR(t) = 1 - VE(t)$, that an unmeasured confounder would need to have with both vaccination status and disease outcome in order to fully explain away a specific

observed vaccine efficacy. Let $\widehat{RR}(t)$ be the estimate of $RR(t)$, and let $UL(t)$ be the upper limit of the 95% confidence interval for $RR(t)$. Then the E-value for $\widehat{RR}(t)$ is given by

$$e(t) = \frac{1 + \sqrt{1 - \widehat{RR}(t)}}{\widehat{RR}(t)},$$

provided that $\widehat{RR}(t) < 1$. In addition, the E-value for $UL(t)$ is computed as 1 if $UL(t) \geq 1$ and as

$$e_{UL}(t) = \frac{1 + \sqrt{1 - UL(t)}}{UL(t)}$$

otherwise. E-values near one suggest weak support for a causal inference, and greater E-values provide increasing evidence for causality.

Suppose, for example, that $\widehat{RR}(t) = 0.50$, with 95% confidence interval $(0.08, 0.75)$. Then $e(t) = 3.4$, meaning that the result of $\widehat{RR}(t)$ being less than one could be explained away by an unmeasured confounder associated with both vaccination status and disease by a risk ratio of 3.4-fold each after accounting for the vector X of measured confounders, but not by a weaker unmeasured confounder. In addition, $e_{UL}(t) = 2.5$, which is the strength of unmeasured confounding at which statistical significance for $VE(t) > 0$ would be lost. These two E-values judge how confident we can be that $VE(t)$ truly exceeds 0, accounting for potential unmeasured confounding due to unblinding as well as for sampling variability.

We can provide a conservative estimate of $VE(t)$ that accounts for potential unmeasured confounding.[8] Let $RR_{UD}(t)$ be the maximum risk ratio for disease when comparing any two categories of the unmeasured confounder U, within either the vaccinated group or the unvaccinated group, conditional on the vector X of observed covariates, and let $RR_{EU}(t)$ be the maximum risk ratio for any specific level of the unmeasured confounder U when comparing the vaccinated and unvaccinated individuals. Define the bias or bounding factor

$$B(t) = \frac{RR_{UD}(t) RR_{EU}(t)}{RR_{UD}(t) + RR_{EU}(t) - 1}.$$

Then a conservative (lower bound) estimate of the vaccine efficacy is given by $1 - \widehat{RR}(t)B(t)$, and a conservative confidence interval is obtained by multiplying the lower and upper limits of the confidence interval for $RR(t)$ by $B(t)$.

References

1. Kalbfleisch JD, Prentice RL (2011). The Statistical Analysis of Failure Time Data, 2nd Ed. John Wiley & Sons.

2. Breslow NE (1972). Discussion of the paper by D. R. Cox. J Roy Stat Soc B, 34: 216-217.

3. Fleming TR, Harrington DP (1991). Counting Processes and Survival Analysis. John Wiley & Sons.

4. Lin DY, Fleming TR, Wei LJ (1994). Confidence bands for survival curves under the proportional hazards model. Biometrika, 81: 73-81.

5. Durham LK, Longini Jr IM, Halloran ME, Clemens JD, Azhar N, Rao M (1998). Estimation of vaccine efficacy in the presence of waning: application to cholera vaccines. American Journal of Epidemiology, 147: 948-959.

6. Ramlau-Hansen H (1983). Smoothing counting process intensities by means of kernel functions. The Annals of Statistics, 11: 453-466.

7. VanderWeele TJ, Ding P (2017). Sensitivity analysis in observational research: introducing the E-value. Ann Int Med, 167: 268-74.

8. Ding P, VanderWeele TJ (2016). Sensitivity analysis without assumptions. Epidemiology, 27: 368.