Clinical Infectious Diseases

INVITED ARTICLE

INNOVATIONS IN DESIGN, EDUCATION AND ANALYSIS (IDEA): Victor De Gruttola and Scott R. Evans, Section Editors

Evaluating the Long-term Efficacy of Coronavirus Disease 2019 (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; and [2]Vaccine and Infectious Disease Division, Fred Hutch, Seattle, Washington, USA

Large-scale deployment of safe and durably effective vaccines can curtail the coronavirus disease-2019 (COVID-19) pandemic. However, the high vaccine efficacy (VE) reported by ongoing phase 3 placebo-controlled clinical trials is based on a median follow-up time of only about 2 months, 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 participants over from the placebo arm to the vaccine arm. Here, we show how to estimate potentially time-varying placebo-controlled VE in this type of staggered vaccination of participants. In addition, we compare the performance of blinded and unblinded crossover designs in estimating long-term VE.

Keywords. booster vaccination; crossover; durability of vaccine efficacy; phase 3 trials; SARS-CoV-2; waning efficacy.

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) [1–12]. Interim results from several large-scale phase 3 randomized, observer-blinded, placebo-controlled clinical trials have demonstrated high vaccine efficacy (VE) [4–6], far exceeding the Food and Drug Administration (FDA) and World Health Organization (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 VE. For example, the recently published results for the Pfizer/BioNTech vaccine BNT162b2 and the Moderna vaccine mRNA-1273 are based on a median follow-up time of approximately 2 months after the second dose; [4–5] therefore, the reported 94%–95% VE pertains only to an average of 2 months post vaccination.

Vaccine effect can wane over time because of declining immunologic memory or changing antigenicity of the pathogen. A vaccination can be followed with booster doses to maintain a protective level of immunity among susceptible individuals, but the nature of the protection over time must be understood so that an effective vaccination and boosting schedule can be determined. Thus, after FDA issues an Emergency Use Authorization (EUA), vaccine sponsors should continue to collect placebo-controlled data on primary endpoints in any ongoing trials for as long as feasible [15].

Although continuing blinded follow-up of the original treatment arms is the ideal way to evaluate long-term 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.

Because participants are vaccinated at different times and community transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2 virus) changes over time, existing statistical methods—which assume that trial participants are vaccinated at the same time or that community transmission is constant over time—do not produce valid estimates of long-term VE in the presence of waning vaccine effect. In fact, participants in COVID-19 vaccine trials are enrolled over several months and thus are randomized to the vaccine and placebo groups at different times; therefore, the timing of vaccination varies among vaccinees, even if the placebo group is maintained throughout the study.

At the time of interim analysis, there were 8 and 162 symptomatic COVID-19 cases in the BNT162b2 vaccine and placebo groups, respectively [4], such that standard methods would estimate VE at 95%. Because of staggered enrollment and time-varying community transmission, this estimate is difficult to interpret if the VE wanes over time. (Adjustment for the participant's follow-up time would not change the estimate since the person years of follow-up are the same between the vaccine and placebo groups. Due to relatively low disease incidence, standard Poisson and Cox models would yield the same estimate [16]).

In this article, we show how to properly assess the durability of VE under staggered enrollment and time-varying community transmission, allowing higher-risk placebo volunteers to get vaccinated earlier than lower-risk ones during the crossover period. Our framework provides unbiased estimation of

Received 20 January 2021; editorial decision 5 March 2021; published online 10 March 2021.
Correspondence: D.- Y. Lin, Department of Biostatistics, University of North Carolina, 3101E McGavran-Greenberg Hall, Chapel Hill, NC 27599-7420, USA (lin@bios.unc.edu).

Clinical Infectious Diseases® 2021;73(10):1927–39

DOI: 10.1093/cid/ciab226

the entire curve of placebo-controlled VE as a function of time elapsed since vaccination, up to the point where most of the placebo volunteers have been vaccinated. We investigate the bias and precision of our approach in estimating long-term VE under various crossover designs, including blinded crossover, in which participants do not know the order of treatments they receive, and unblinded crossover, in which participants are notified of their randomization assignments at the time of crossover. We also 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: participants are screened and randomly assigned to vaccine or placebo over a 4-month period, the vaccine is granted EUA in the 5th month on the basis of interim results, and crossover occurs over the next 5 months. This design provides information about VE for 10 months. We aim to estimate time-varying VE under any crossover design, up to the point when there are very few placebo participants left.

The endpoint of interest is time to symptomatic COVID-19 disease. We allow the risk of disease to vary over the calendar time and to depend on baseline risk factors, such as age, sex, ethnicity, race, occupation, and underlying health conditions; we allow the effect of vaccine on disease occurrence to depend on the time elapsed since vaccination.

We consider 2 definitions of time-varying VE: (1) day-t VE is the percentage reduction in the hazard rate or instantaneous risk of disease at day t for those who were vaccinated t days ago compared with those who have not been vaccinated; and (2) t-day VE is the percentage reduction in the attack rate or cumulative incidence of disease over the t-day period for those who were vaccinated at the start of the period, compared with those who were unvaccinated throughout the period. We denote these 2 VE measures by $VE_h(t)$ and $VE_a(t)$, where h and a stand for hazard rate and attack rate, respectively. These 2 definitions are equivalent when the effect of vaccine is constant over time. If the effect of vaccine wanes over time, then VE_a is larger than VE_h. It is also of interest to consider VE_a (ie, percentage reduction in the attack rate) over successive time periods, say every quarter.

In Supplementary Appendix 1, we formulate the above concepts through an adaptation of the well-known Cox regression model [17, 18], in which each participant's time to disease occurrence is measured from a common origin, namely the start of the clinical trial, and the hazard ratio of vaccine versus placebo depends on the time elapsed since vaccination. We formally define VE_a as 1 minus the time-averaged hazard ratio, which is approximately the ratio of the cumulative incidence. We derive the maximum likelihood estimator for VE_a as a function of time elapsed since vaccination. We show that the estimator is approximately unbiased and normally distributed, with a variance that can be estimated analytically, enabling one to construct valid confidence intervals for the VE_a curve. In addition, we propose a method to estimate VE_h by kernel-smoothing the estimated VE_a curve. Finally, we show how to estimate VE_a over successive time periods.

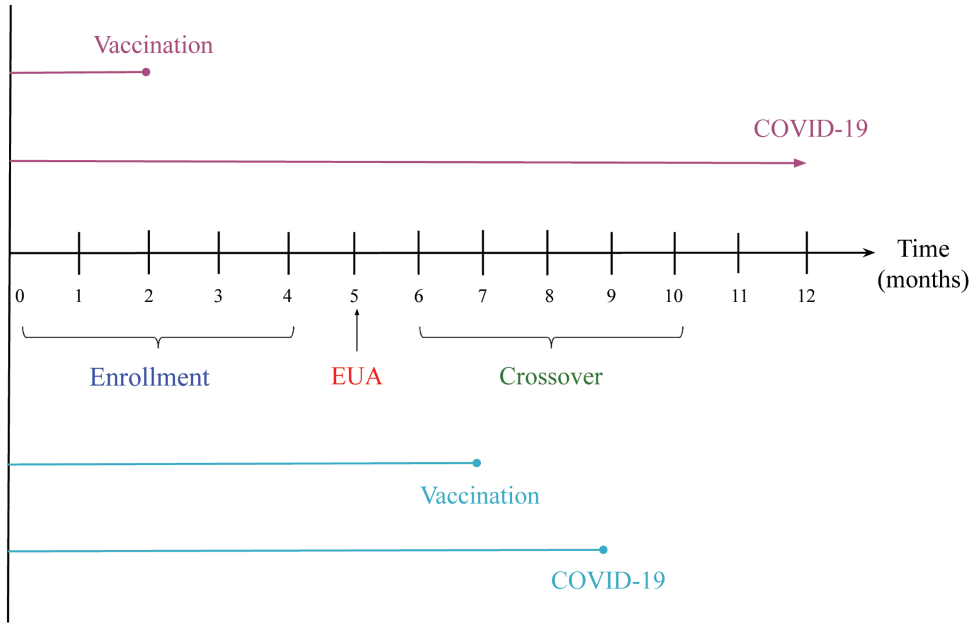

Figure 1. A phase 3 COVID-19 vaccine trial. Participants are enrolled over a 4-month period, EUA is issued at month 5, and crossover occurs over months 6–10. The top 2 lines represent a participant who is vaccinated at month 2 and does not develop symptomatic COVID-19 during the trial; the bottom lines represent a participant who is vaccinated at month 7 and develops symptomatic COVID-19 at month 9.

RESULTS

We conducted a set 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 vaccine rollout and other factors, we let the disease risk increase over the first 7 months and decrease afterward. We chose 3 combinations of 5-month VE and 10-month VE: (a) VE_a(5 months) = VE_a(10 months) = 95%; (b) VE_a(5 months) = 85%, VE_a(10 months) = 75%; and (c) VE_a(5 months) = 70%, VE_a(10 months) = 50%.

We considered the statistically optimal design of keeping all participants on their original treatment assignments until the end of the trial. We refer to this design as Plan A and regard it as a benchmark. We also considered 3 blinded crossover designs:

B. Crossover starts at month 6, 7, 8, 9, or 10 for participants with priority tier of 1, 2, 3, 4, or 5, respectively, with each participant's waiting time for the clinic visit following the exponential distribution with mean of 0.5 month.
C. 20% of participants follow Plan A, and the rest follow Plan B.
D. Crossover starts at month 6 for all participants, with the waiting time following the exponential distribution with mean of 0.5 month.

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. Under Plan D, all placebo recipients are vaccinated quickly without any priority tiering. With blinded crossover, placebo participants receive the vaccine and vaccine participants receive the placebo at the point of crossover; none of the participants are aware of the order of their treatments. All participants are followed until the time of analysis, which is 10.5 months since trial initiation. The designs of these simulation studies are detailed in Supplementary Appendix 2.

The results for the estimation of VE_a based on 10 000 simulated datasets are summarized in Table 1 and Figure 2. The proposed method yields virtually unbiased estimates of the VE_a curves over the 10-month period for Plans A–C in all 3 scenarios of long-term VE; it also yields accurate variance estimates, such that the confidence intervals have correct coverage probabilities. When VE is constant over time, the standard errors for the estimates of VE_a under Plans B and C are slightly lower

Table 1. Estimation of Time-Varying VE_a by Proposed and Standard Methods Under No Crossover (A) and Three Blinded Crossover Plans (B–D)

			True	Proposed method				Standard	
Scenario	Plan	Time	VE_a	Mean	SE	SEE	CP	Mean	SE
(a)	A	5 m	95%	95.0%	0.8%	0.8%	95.2%		
		10 m	95%	95.0%	0.8%	0.8%	94.5%	95.0%	0.6%
	B	5 m	95%	95.0%	0.7%	0.7%	95.1%		
		10 m	95%	95.1%	0.8%	0.7%	94.5%	95.0%	0.6%
	C	5 m	95%	95.0%	0.7%	0.7%	95.2%		
		10 m	95%	95.0%	0.7%	0.7%	94.5%	95.0%	0.6%
	D	5 m	95%	95.0%	0.8%	0.8%	94.8%		
		10 m	95%	95.2%	0.9%	0.9%	93.6%	95.0%	0.7%
(b)	A	5 m	85%	84.9%	1.4%	1.5%	95.4%		
		10 m	75%	74.9%	2.2%	2.2%	94.9%	78.5%	1.4%
	B	5 m	85%	85.0%	1.3%	1.3%	95.1%		
		10 m	75%	75.0%	2.6%	2.6%	94.9%	81.6%	1.3%
	C	5 m	85%	85.0%	1.3%	1.3%	95.7%		
		10 m	75%	75.0%	2.4%	2.4%	94.8%	80.9%	1.3%
	D	5 m	85%	85.0%	1.5%	1.5%	94.7%		
		10 m	75%	74.9%	3.4%	3.4%	94.6%	84.6%	1.4%
(c)	A	5 m	70%	69.9%	2.2%	2.2%	95.1%		
		10 m	50%	49.9%	3.4%	3.3%	94.5%	57.2%	2.2%
	B	5 m	70%	69.9%	2.0%	2.0%	95.0%		
		10 m	50%	50.0%	4.1%	4.1%	95.5%	64.0%	2.0%
	C	5 m	70%	70.0%	2.0%	2.0%	95.0%		
		10 m	50%	49.9%	3.8%	3.8%	94.7%	62.5%	2.0%
	D	5 m	70%	70.0%	2.3%	2.3%	94.8%		
		10 m	50%	49.9%	5.1%	5.1%	94.9%	69.5%	2.2%

Abbreviations: CP, coverage probability of 95% confidence interval; m, months; Mean and SE, mean and standard error of the estimator for VE_a; SEE, mean of standard error estimator; VE, vaccine efficacy.

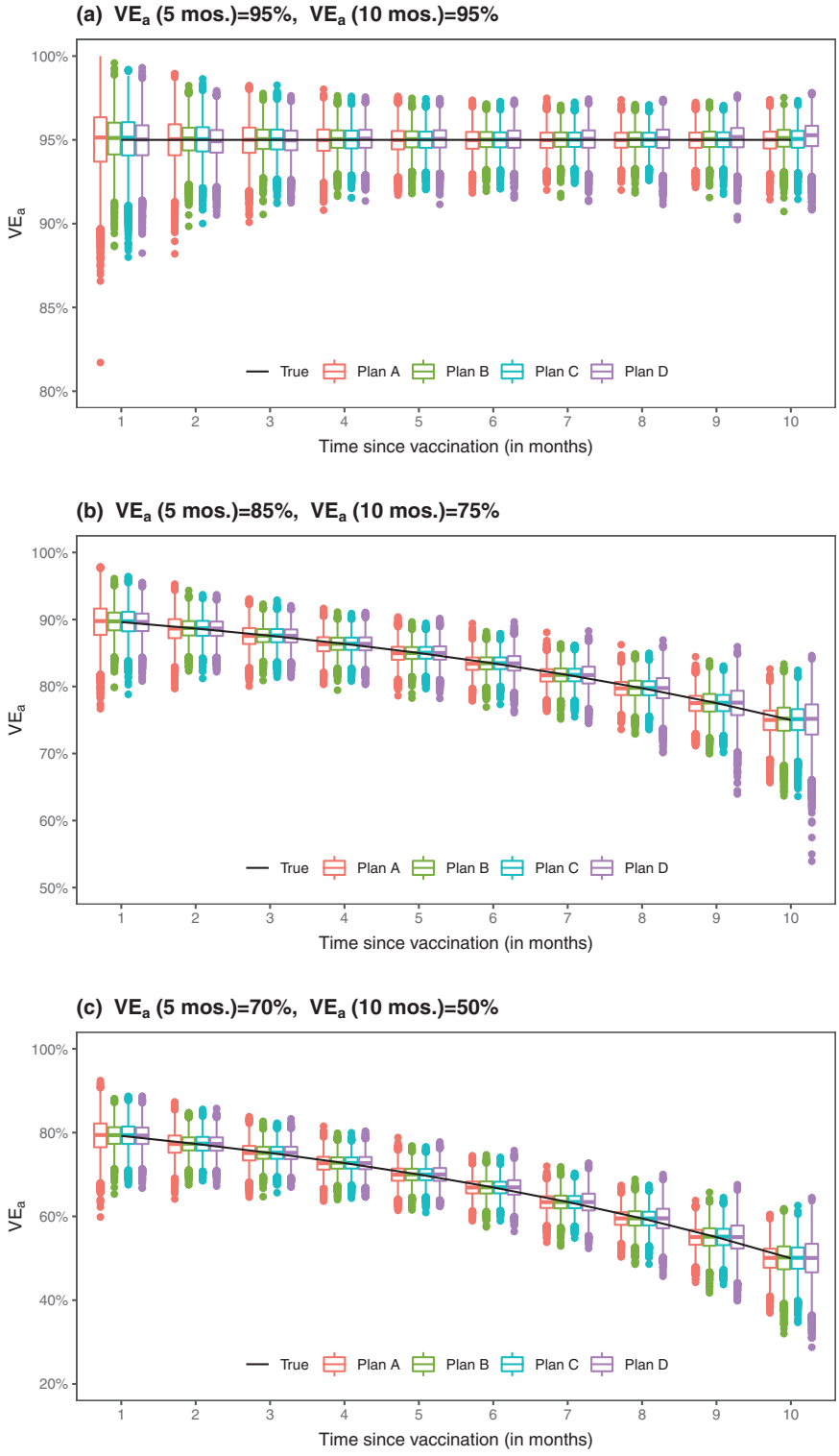

Figure 2. Proposed estimates of VE_a under no crossover and 3 blinded crossover plans.

than those of Plan A. When VE wanes over time, the standard errors for the estimates of 5-month VE_a under Plans B and C are also slightly lower than those of Plan A; however, the standard errors for the estimates of 10-month VE_a 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 VE_a may be slightly biased, with higher standard errors than under Plans A–C; 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 standard Cox regression [17, 18], using vaccine status as a potentially time-dependent covariate with a constant hazard ratio. Because it estimates an overall VE among individuals who have been vaccinated for different amounts of time, this method overestimates long-term VE when VE decreases over time, although the estimation is unbiased when VE is constant over time. Even though the results for standard Cox regression are shown in the "10-month" row of the table, the estimates do not pertain to the hazard ratio at month 10 or to the average hazard ratio over the first 10 months. In scenario (c), where 10-month VE_a equals 50% and month-10 VE_h equals –3.8%, standard Cox regression yields mean VE estimates of 57.2%, 64.0%, 62.5%, and 69.5% under Plans A, B, C, and D, respectively; the poor performance under Plan A highlights the fact that standard Cox regression does not properly capture waning VE, even when all placebo recipients remain on their original assignments until the end of the trial.

We conducted a second set of simulation studies by considering 4 unblinded crossover designs:

B'. Crossover occurs at month 6.5, 7.5, 8.5, 9.5, or 10.5 for participants with priority tier of 1, 2, 3, 4, or 5, respectively.
C'. 20% of participants follow Plan A, and the rest follow Plan B'.
D'. Crossover occurs at month 6.5 for all participants.
D." Crossover starts at month 6 for all participants, with the waiting time following the exponential distribution with mean of 0.5 month.

With unblinded crossover, participants are notified of their original treatment assignments at the point of crossover, and placebo recipients are vaccinated soon after. In Plan B', everyone in a given priority-tier group is notified of their original treatment assignment on the same day. In Plan D', all participants are notified of their original treatment assignments on the same day. In Plan D," participants are notified of their randomization assignments without priority tiering; the timing of crossover is the same as that of Plan D. Plan D" was meant to mimic the crossover that has been occurring in the 2 mRNA vaccine trials, where participants have been unblinded gradually.

Because vaccine recipients may engage in riskier behavior upon unblinding and placebo recipients may also change their behavior upon unblinding (in a manner that likely differs from vaccine recipients), 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. This strategy avoids bias due to behavioral confounding.

The results for the estimation of VE_a based on 10 000 simulated datasets are summarized in Table 2 and Figure 3. The proposed method yields unbiased estimates of the VE_a curves under Plans B' and C', but the standard errors for estimates of long-term VE are markedly higher than under Plans B and C; the standard errors estimates are accurate only up to month 8. Under Plan D', where all placebo recipients are vaccinated at month 6.5, the estimates of 5-month VE_a are unbiased; however, the estimates of 8-month or 10-month VE_a are biased, especially when VE wanes over time. (Using the unblinded data would not reduce the bias in estimating long-term VE because there are no placebo participants beyond month 6.5.) Of note, the bias of standard Cox regression in estimating waning VE is more severe under unblinded crossover than under blinded crossover. This result reflects the fact that standard statistical methods do not provide valid estimates of waning VE even in the absence of crossover—because time to disease is censored at time of unblinding, there is no actual crossover of treatments in this analysis.

The results under Plan D" are encouraging: the estimates of the VE_a curves have little bias, and the standard errors are accurately estimated, such that confidence intervals have proper coverage probabilities—at least for the first 8 months. The timing of crossover is the same between Plans D and D"; however, we disregarded the follow-up data collected after unblinding, such that the precision of estimates is lower under Plan D" than under Plan D. Although the mean time to crossover under Plan D" is the same as that of Plan D', crossover spreads over a longer period under Plan D," making it possible to estimate long-term VE.

Figures 4 and 5 show the results for estimating VE_h from the first and second sets of simulation studies, respectively. Under Plans A–D, the estimates of the VE_h curves are virtually unbiased, at least up to month 9. Under Plans B' and C', the VE_h estimates are nearly unbiased up to month 8. Under Plan D', VE_h is estimated well until month 5. Under Plan D," the estimates of the VE_h curves are virtually unbiased, except for the last 2 months in scenario (a).

Figure 6 displays the estimation results for VE_a produced by the proposed methods in 1 of the trials simulated with 75% 10-month VE_a. The estimates of VE_a are close to the truth, and the 95% confidence intervals cover the truth up to the end of crossover. In terms of estimating long-term VE_a, 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 considerably better than Plans B', C', and D', respectively; and Plan D" is much better than Plan D'.

Figure 7 presents the analysis results for VE_h from the same trial. The estimates are close to the truth up to the point where there are still a few placebo participants under follow-up. The estimates are more bumpy under unblinded than blinded crossover. Comparing Figures 6 and 7 shows that VE_h is much lower than VE_a in the presence of waning vaccine effect.

We also evaluated our method for estimating VE_a over successive time periods. We found that the method performs well

Table 2. Estimation of Time-Varying VE_a by Proposed and Standard Methods Under 4 Unblinded Crossover Plans

			True	Proposed method				Standard	
Scenario	Plan	Time	VE_a	Mean	SE	SEE	CP	Mean	SE
(a)	B′	5 m	95%	95.0%	0.9%	0.9%	95.3%		
		8 m	95%	95.0%	0.9%	0.9%	94.0%		
		10 m	95%	95.1%	2.0%	1.3%	85.0%	95.0%	0.8%
	C′	5 m	95%	95.0%	0.9%	0.9%	95.1%		
		8 m	95%	95.0%	0.9%	0.9%	94.2%		
		10 m	95%	94.9%	7.1%	1.5%	83.9%	95.0%	0.8%
	D′	5 m	95%	95.0%	1.0%	1.0%	94.6%		
		8 m	95%	96.3%	1.3%	0.9%	65.1%		
		10 m	95%	97.0%	1.1%	0.7%	29.7%	95.0%	1.0%
	D″	5 m	95%	95.0%	0.8%	0.8%	95.3%		
		8 m	95%	95.2%	0.8%	0.8%	94.1%		
		10 m	95%	95.5%	0.8%	0.8%	88.8%	95.0%	0.8%
(b)	B′	5 m	85%	85.0%	1.6%	1.6%	95.6%		
		8 m	80%	79.7%	2.3%	2.3%	95.3%		
		10 m	75%	75.8%	7.6%	5.2%	83.1%	82.5%	1.5%
	C′	5 m	85%	85.0%	1.6%	1.6%	94.9%		
		8 m	80%	79.7%	2.3%	2.3%	94.7%		
		10 m	75%	75.0%	8.6%	5.7%	84.9%	82.4%	1.5%
	D′	5 m	85%	85.0%	1.9%	1.9%	94.8%		
		8 m	80%	87.5%	3.1%	2.1%	16.4%		
		10 m	75%	90.0%	2.5%	1.7%	2.6%	85.5%	1.7%
	D″	5 m	85%	85.0%	1.5%	1.5%	95.4%		
		8 m	80%	79.7%	2.5%	2.5%	94.9%		
		10 m	75%	75.0%	4.4%	4.2%	93.8%	83.5%	1.5%
(c)	B′	5 m	70%	70.0%	2.3%	2.3%	95.6%		
		8 m	60%	59.4%	3.4%	3.5%	95.2%		
		10 m	50%	51.1%	16.4%	8.6%	85.8%	64.9%	2.3%
	C′	5 m	70%	69.9%	2.3%	2.3%	95.0%		
		8 m	60%	59.3%	3.4%	3.4%	95.1%		
		10 m	50%	49.8%	21.3%	9.1%	87.2%	64.6%	2.3%
	D′	5 m	70%	69.9%	2.8%	2.8%	95.0%		
		8 m	60%	75.0%	5.3%	3.2%	7.5%		
		10 m	50%	80.0%	4.2%	2.6%	1.0%	71.0%	2.5%
	D″	5 m	70%	69.9%	2.3%	2.3%	94.9%		
		8 m	60%	59.3%	4.0%	3.9%	94.6%		
		10 m	50%	49.7%	6.8%	6.6%	94.6%	67.8%	2.3%

Abbreviations: CP, coverage probability of 95% confidence interval; m, months; Mean and SE, mean and standard error of estimator for VE_a; SEE, mean of standard error estimator; VE, vaccine efficacy.

when the time interval is at least one-month wide. For example, Table 3 summarizes the results for estimating monthly VE_a under Plan D" in scenario (b). The VE_a estimates are virtually unbiased, the standard errors are accurately estimated, and the confidence intervals have proper coverage probabilities. Table 4 shows the analysis results on bimonthly VE_a for the trial used in Figures 6–7. Bimonthly VE_a is estimated with higher precision than monthly VE_a.

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 volunteers to be vaccinated in a timely manner while still making it possible to assess long-term vaccine safety and efficacy. As our simulation studies have shown, standard Cox regression with a constant hazard ratio seriously overestimates long-term VE in the presence of waning vaccine effect. We have developed a valid and efficient approach to evaluate the effect of a COVID-19 vaccine that potentially wanes over time. The estimated curve of time-varying VE can be used to determine when a booster vaccination is needed to sustain protection; this

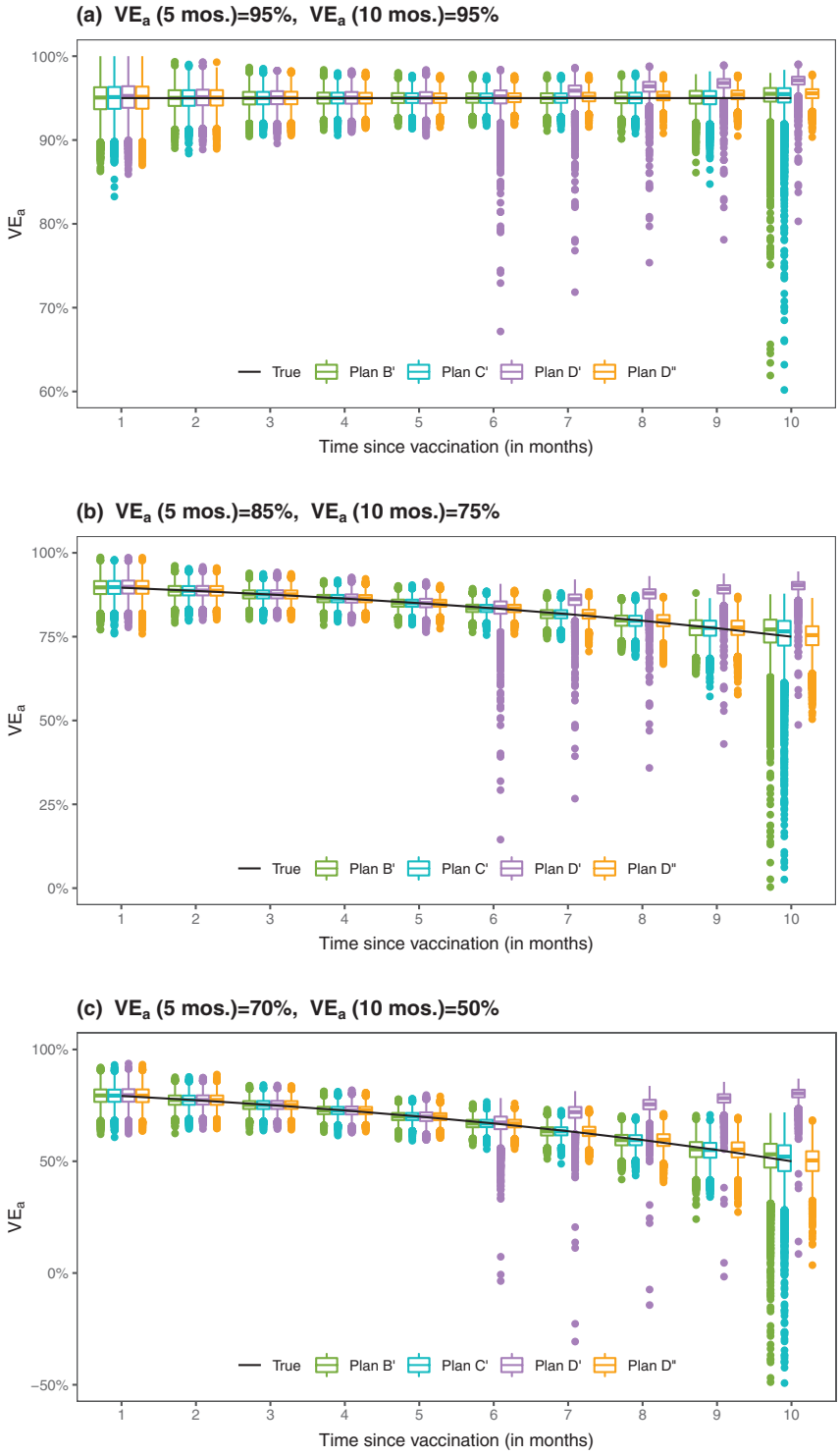

Figure 3. Proposed estimates of VE_a under 4 unblinded crossover plans.

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 participants not knowing their treatment assignments, even after crossover. It is advantageous, when possible, to implement crossover on a rolling basis rather than instantaneously since time-varying VE can be estimated (without adding assumptions) only up to the point where there are still a few placebo recipients under follow-up. Indeed, rolling crossover is even more important than blinding for the express purpose of assessing long-term VE without imposing additional assumptions.

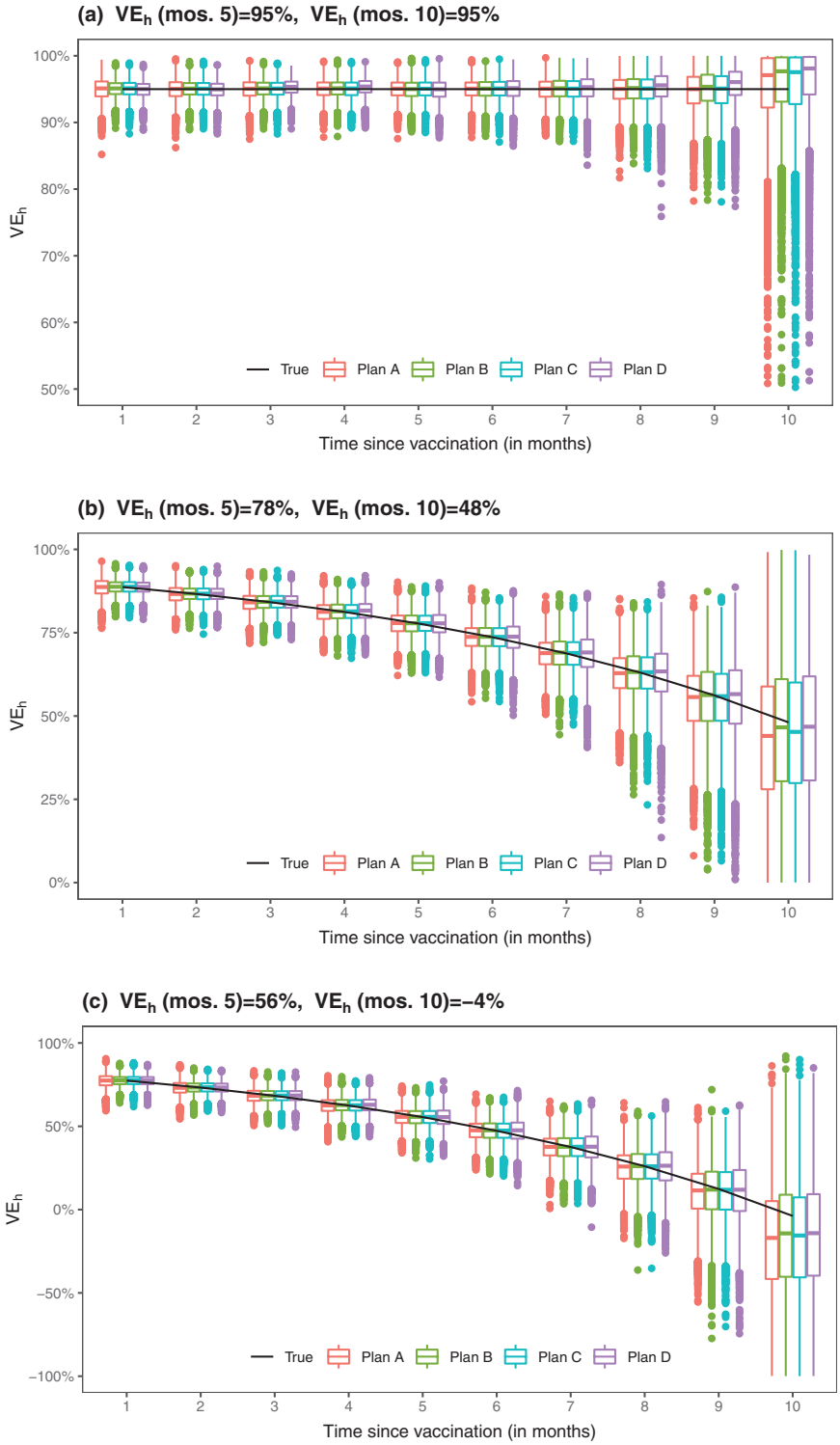

Figure 4. Proposed estimates of VE_h under no crossover and 3 blinded crossover plans.

Of course, 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 VE. 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 VE. We may estimate VE twice, once with all follow-up data, and once with only blinded follow-up data, and compare the 2 sets of results.

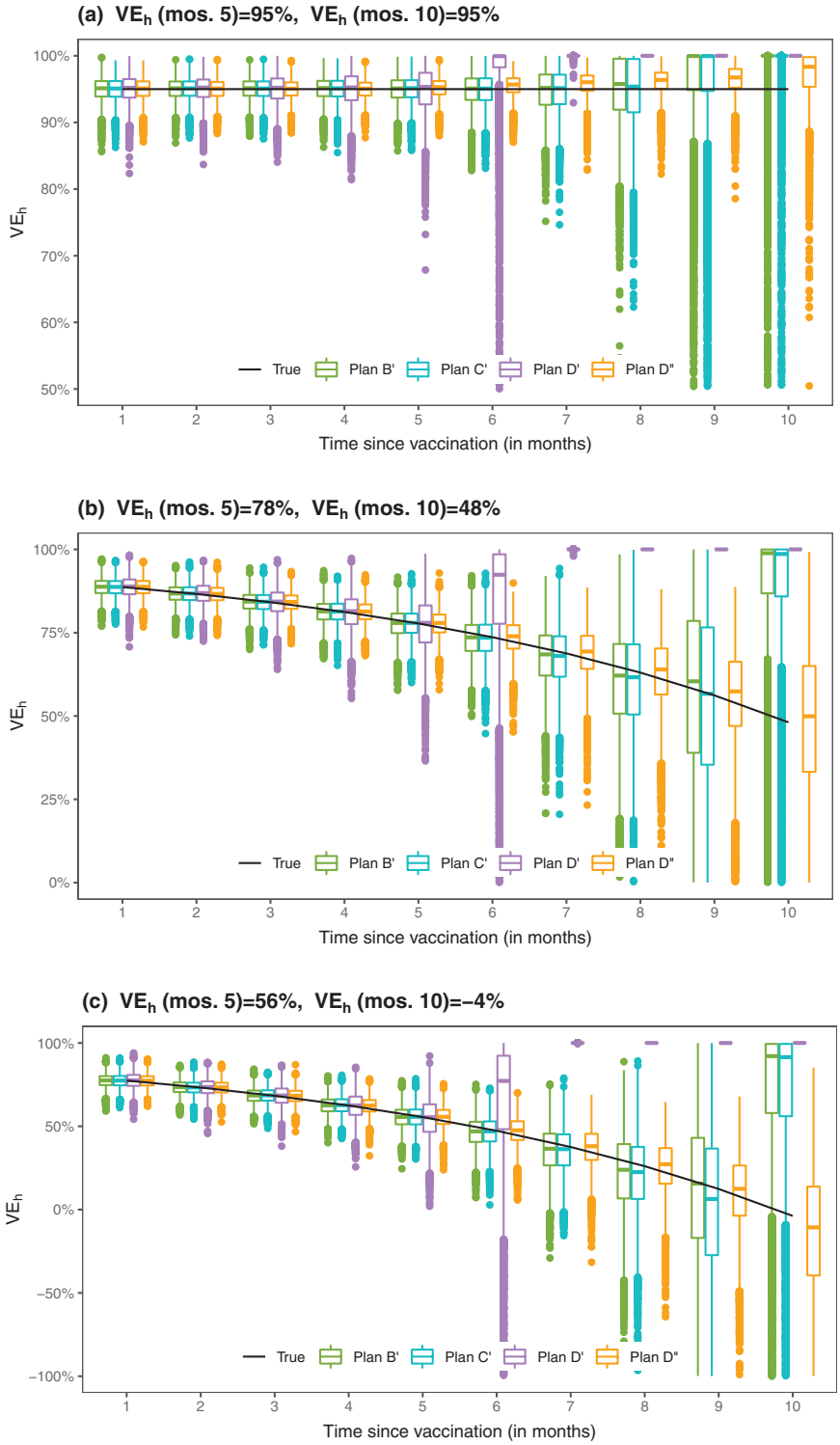

Figure 5. Proposed estimates of VE_h under 4 unblinded crossover plans.

Alternatively, our methods can be applied to all follow-up data, followed by a sensitivity analysis to assess the robustness of the results to potential unmeasured confounding caused by unblinding of trial participants. In Supplementary Appendix 3, we show how to apply a best-practice general methodology in epidemiological research [19–20] to perform this sensitivity analysis. Using this methodology, we can assess how strong unmeasured confounding due to unblinding would need to be in order to fully explain away the observed VE. We can also provide a conservative estimate of VE that accounts for unmeasured confounding.

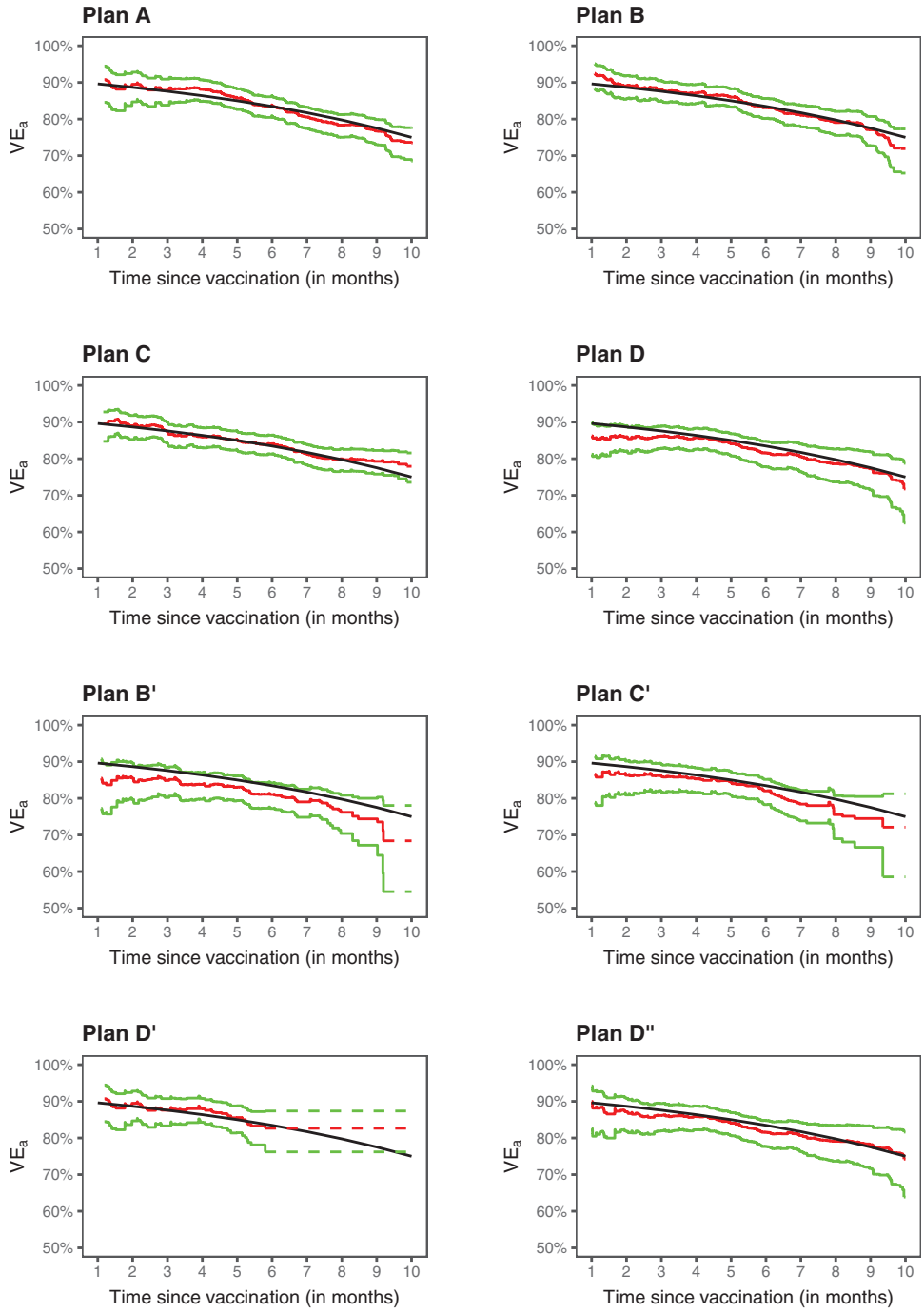

Figure 6. Estimation of VE_a in a clinical trial with no crossover (A), 3 blinded crossover (B–D), and 4 unblinded crossover (B′–D″): the black curve pertains to the true value, the red curve to the proposed estimate, and the green curves to the 95% confidence intervals.

Recently, Follmann et al [21] advocated blinded crossover and continued follow-up of trial participants to assess vaccine durability and potential delayed enhancement of disease. They estimated an overall VE for the original vaccine arm in the postcrossover period (after all placebo volunteers were vaccinated) by imputing the case count for a counter-factual placebo group under certain assumptions. We address the issue of VE durability in a different way, using observed data to estimate the entire curve of VE 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.

The methods we have described assess overall VE against all viral variants of COVID-19. The results may be difficult to interpret if the distribution of viral variants changes during the

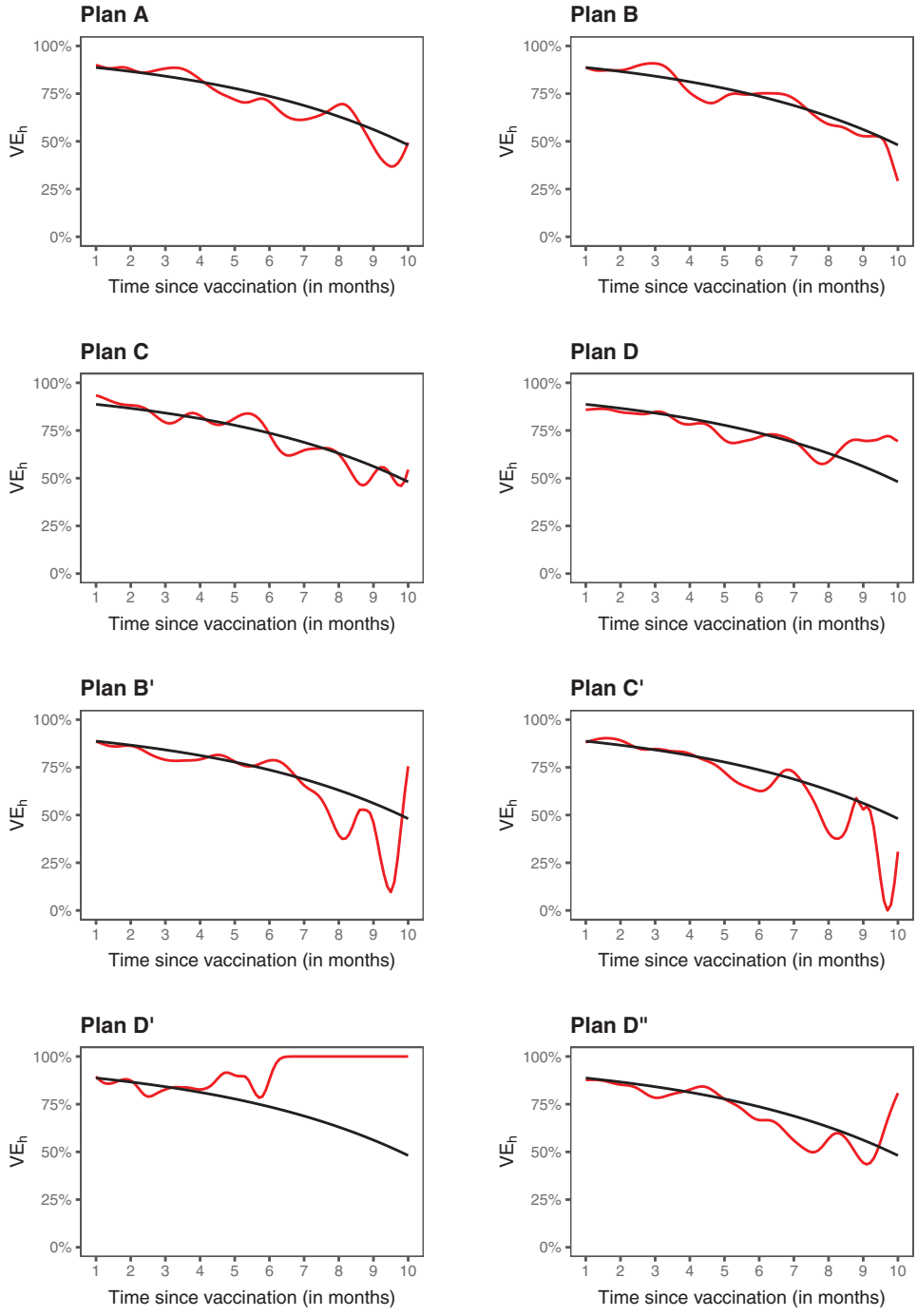

Figure 7. Estimation of VE_h in a clinical trial with no crossover (A), 3 blinded crossover (B–D), and 4 unblinded crossover (B′–D″): the black curve pertains to the true value, and the red curve to the proposed estimate.

trial, because waning VE might be caused by increasing prevalence of resistant variants, while VE against specific variants remains constant. It is possible to address this challenge if SARS-CoV-2 sequences from all COVID-19 cases are measured. Our methods can then be applied to the subset of COVID-19 cases caused by each specific variant or set of variants. Conducting this analysis for each spectrum of viral genotypes provides interpretable results on the durability of VE.

Although we have focused on VE in the entire study population, we can also estimate VE for various subgroups, such as age group, sex, and race/ethnic group, by applying our methods to a subset of participants. Because we allow the effect of vaccination to vary over time in an arbitrary manner, however, our estimates of long-term VE may be unstable if there are only a small number of cases in a subgroup. To alleviate this problem, we may formulate the time-varying hazard ratio through a

Table 3. Estimation of Monthly VE_a Under Unblinded Crossover Plan D"

Months	True VE_a	Bias	SE	SEE	CP
0–1	89.6%	0.0%	2.9%	2.9%	95.4%
1–2	87.7%	-0.1%	3.0%	3.0%	95.6%
2–3	85.4%	0.0%	3.0%	3.1%	95.8%
3–4	82.7%	-0.1%	3.3%	3.3%	95.3%
4–5	79.5%	-0.1%	3.8%	3.8%	95.3%
5–6	75.8%	-0.2%	4.8%	4.8%	95.5%
6–7	71.3%	-0.1%	6.6%	6.5%	95.1%
7–8	66.0%	0.0%	9.2%	9.0%	95.0%
8–9	59.7%	0.1%	12.8%	12.6%	95.2%
9–10	52.3%	0.3%	19.4%	18.9%	95.2%

Note: Bias and SE denote the bias and standard error of the estimator for VE_a, SEE denotes the mean of the standard error estimator, and CP denotes the coverage probability of the 95% confidence interval.

parametric (eg, log-linear) function, which is then allowed to interact with subgroups.

We have targeted VE over the first 10 months for several reasons. First, it is unlikely that there will be any placebo volunteers beyond month 10 in the ongoing COVID-19 vaccine trials, and follow-up data in the absence of a placebo arm do not provide direct information about VE durability. Second, estimates of long-term VE will become more uncertain as community transmission decreases.

The primary endpoints for most phase 3 trials are symptomatic COVID-19 cases at least one or two weeks after the second dose. Our methods can be used to estimate the entire VE curve starting on the day of the first dose (or the day of injecting a single-dose vaccine).

Although we have framed the discussion in the context of randomized, placebo-controlled phase 3 trials, our methods do not require the designation of vaccine and placebo groups; instead, everyone is considered a potential vaccinee, with the only distinction being *when* they are vaccinated. This general framework can be applied to surveillance data in order to estimate the risk of disease as a function of time elapsed since vaccination in the "real world." With the large volume of surveillance data, we can estimate the effectiveness of vaccination in various subpopulations and against different viral variants, as well as the duration of any protective effect. Causal interpretations of results, however, require the timing of vaccination to depend only on observed covariates. If it is impossible to measure the key timing factors or inappropriate to include them as covariates in the model, then the aforementioned sensitivity analysis would be warranted.

Table 4. Estimation of Bimonthly VE_a in a Clinical Trial

Months	Estimate	SE	95% Confidence Interval
0–2	88.7%	2.1%	(83.7%, 92.2%)
2–4	88.1%	2.0%	(83.6%, 91.4%)
4–6	80.4%	2.9%	(74.0%, 85.3%)
6–8	78.5%	4.9%	(66.5%, 86.2%)
8–10	67.4%	9.5%	(42.2%, 81.6%)

We have implemented the methods described in this article in an R package, which is available at https://dlin.web.unc.edu/software/dove/.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Nonstandard Abbreviations. COVID-19, coronavirus disease-2019; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; VE, vaccine efficacy; FDA, Food and Drug Administration; WHO, World Health Organization; EUA, Emergency Use Authorization.

Notes

Acknowledgments. The authors are grateful to Yu Gu and Bridget I. Lin for assistance, to Thomas Fleming, David Harrington, and Ross Prentice for helpful discussions.

Financial support. This work was supported by the National Institutes of Health grants R01 AI029168, R01 GM124104, P01 CA142538, and UM1 AI068635.

Potential conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Krause P, Fleming TR, Longini I, et al. COVID-19 vaccine trials should seek worthwhile efficacy. Lancet **2020**; 396:741–3.
2. Cohen MS, Corey L. Combination prevention for COVID-19. Science **2020**; 368:551.
3. Corey L, Mascola JR, Fauci AS, Collins FS. A strategic approach to COVID-19 vaccine R&D. Science **2020**; 368:948–50.
4. Polack FP, Thomas SJ, Kitchin N, et al; C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med **2020**; 383:2603–15.
5. Baden LR, El Sahly HM, Essink B, et al; COVE Study Group. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med **2021**; 384:403–16.
6. Voysey M, Clemens SAC, Madhi SA, et al; Oxford COVID Vaccine Trial Group. 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 **2021**; 397:99–111.
7. Xia S, Duan K, Zhang Y, et al. 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 **2020**; 324:951–60.
8. Jackson LA, Anderson EJ, Rouphael NG, et al. An mRNA vaccine against SARS-CoV-2 – Preliminary report. New Eng J Med **2020**; 383:1920–31.
9. Folegatti PM, Ewer KJ, Aley PK, et al; Oxford COVID Vaccine Trial Group. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase ½, single-blind, randomised controlled trial. Lancet **2020**; 396:467–78.
10. Keech C, Albert G, Cho I, et al. Phase 1–2 trial of a SARS-CoV-2 recombinant spike protein nanoparticle vaccine. New Eng J Med **2020**; 383:2320–32.
11. The World Health Organization. Draft landscape of COVID-19 candidate vaccines. **2020**. Available at: https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate- vaccines. Accessed 8 December 2020.
12. The World Health Organization. An international randomised trial of candidate vaccines against COVID-19. **2020**. Available at: https://www.who.int/publications-detail/an-international-randomised-trial-of-candidate-vaccines-against-covid-19. Accessed 24 March 2021.
13. U.S. Food and Drug Administration. Development and licensure of vaccines to prevent COVID-19: guidance for industry. **2020**. Available at: https://www.fda.gov/media/139638/download. Accessed 24 March 2021.

14. The World Health Organization. WHO target product profiles for COVID-19 vaccines. **2020**. Available at: https://www.who.int/who-documents-detail/who-target-productprofiles- for-covid-19-vaccines. Accessed 24 March 2021.
15. U.S. Food and Drug Administration. Emergency use authorization for vaccines to prevent COVID-19: guidance for industry. **2021**. Available at: https://www.fda.gov/media/142749/download. Accessed 24 March 2021.
16. Lin DY, Zeng D, Mehrotra DV, Corey L, Gilbert PB. Evaluating the efficacy of COVID-19 vaccines. Clin Infect Dis **2021**; in press doi: 10.1093/cid/ciaa1863
17. Cox DR. Regression models and life-tables. J Roy Stat Soc B **1972**; 34:187–202.
18. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data. 2nd edn. Hoboken: John Wiley & Sons, **2011**.
19. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. Ann Int Med **2017**; 167:268–74.
20. Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. Epidemiology **2016**; 27:368–77.
21. Follmann D, Fintzi J, Fay MP, et al. Assessing durability of vaccine effect fol- lowing blinded crossover in COVID-19 vaccine efficacy trials. **2020**. Available at: https://www.medrxiv.org/content/10.1101/2020.12.14.20248137v1. Accessed 24 March 2021.