Supplementary materials

Illness onset hazard rates and ratio

We provide an example where the background infection hazard rate decreases over time. In a ring vaccination trial, after each ring is identified and randomized, a decrease in the background hazard rate of infection could occur because of increased awareness in the community, modified behavior, or an exhaustion of susceptibles in the population. FigureS8 depicts the illness onset hazard rates and apparent VE as a function of time in the standard setting with linearly decreasing background hazard rate of infection. The impact on hazard rates can be compared to Figure 2. It is noteworthy that the bias (difference between apparent $VE_T(t)$ and true VE_0) at time t is the same regardless of background hazard because it cancels across trial arms.

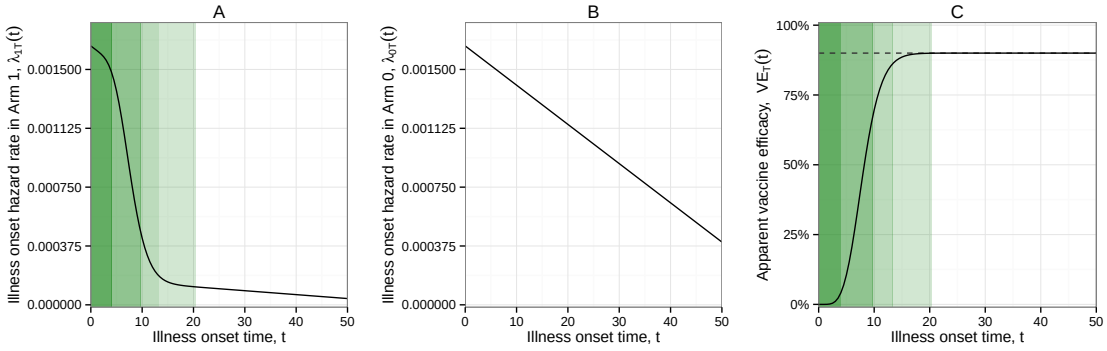

Figure S8: Arm 1 (intervention arm) vaccinated on day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($s_1 = \infty$). Background infection hazard rate decreasing linearly from $\lambda_W(0) = 0.0015$ to $\lambda_W(60) = 0$. Standard setting. (A) Illness onset hazard rate in Arm 1 as a function of time. (B) Illness onset hazard rate in Arm 0 as a function of time. (C) Apparent VE as a function of time.

In Figure S9, we provide an example where the background infection hazard rate decreases over time and delayed vaccination is used. Note that Panel C depicting the apparent vaccine efficacy in this setting is identical to that in Panel C in Figure 3, for which the background infection hazard rate is constant over time; this is because at any given time t the background infection hazard rate cancels in the ratio.

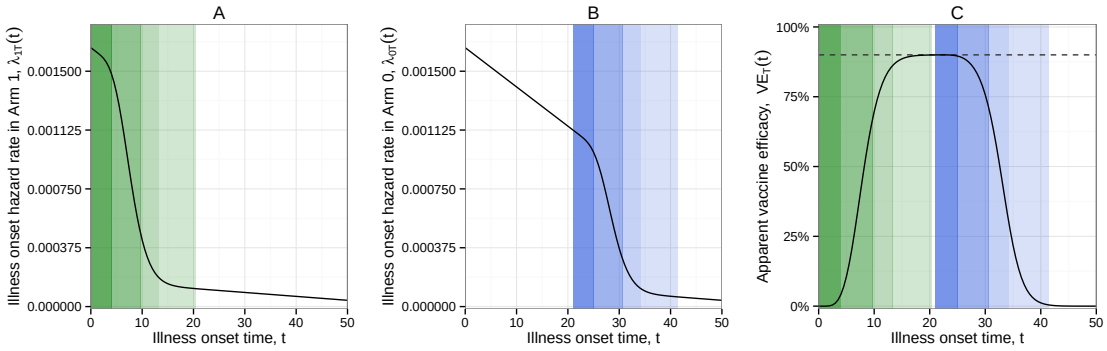

Figure S9: Arm 1 (intervention arm) vaccinated on day $s_1 = 0$; Arm 0 (comparator arm) vaccinated on day $s_1 = 21$ ($b = 21$). Background infection hazard rate decreasing linearly from $\lambda_W(0) = 0.0015$ to $\lambda_W(60) = 0$. Standard setting. (A) Illness onset hazard rate in Arm 1 as a function of time. (B) Illness onset hazard rate in Arm 0 as a function of time. (C) Apparent VE as a function of time.

Trial analysis framework

We use simulations to evaluate our closed-form approximations of apparent vaccine efficacy and power for trial designs. Trials are simulated with individual randomization with n participants per arm. Individuals within each arm are vaccinated on the same day. For unvaccinated participants, infections are drawn from an exponential distribution with infection hazard rate equal to the background infection hazard rate. For vaccinated participants, their hazard rate is multiplicatively reduced by the vaccine efficacy (leaky vaccine effect). Given infection, incubation periods are drawn from a gamma distribution yielding observed illness onset times. Individuals with illness onset times prior to the analysis period are excluded. A Cox proportional hazards model is fit for the trial analysis period, with time rescaled so the first day of the analysis period is day 0. Individuals with illness onset times after the analysis period are censored. We estimate vaccine efficacy as one minus the estimated hazard ratio from the Cox model, and we retain the p-value to be compared to $\alpha = 0.05$. 25,000 simulations are run for each scenario (row within a table) in R [R Core Team, 2015] using the survival package [Therneau, 2015].

Results for $n = 1000$ are shown in Table S1. Results for $n = 500$ and $n = 5000$ are shown in Tables S2 and S3, respectively. Table S4 provides results for $n = 1000$ with a 21 day vaccination delay.

The approximation for the point estimate works best when there are more events, either due to higher background infection hazard rate λ_W or due to a larger sample size per arm n. In general, the simulated mean VE approaches the approximated value from below as the number of events increases; thus, the simulated mean is always below the approximated mean. When the sample size is $n = 1000$ per arm, the greatest discrepancy occurs when $VE_0 = 0\%$ (difference of -3.9%). The smallest discrepancy is when

$VE_0 = 90\%$. Generally, the simulated and approximated VE values are within 1-2% (absolute difference).

Approximated and simulated power track well together, though approximated power tends to slightly overestimate power. Similar to bias, the approximation for power works best when there are more events, either due to higher background infection hazard rate λ_W or due to a larger sample size per arm n. Furthermore, when vaccine efficacy is high ($VE_0 = 90\%$), there are greater discrepancies because fewer events in the immediately vaccinated arm results yields more instability in testing. An extreme scenario is provided in Table S2 in which fewer than 5 events are expected in the immediate arm; in that case, the power approximation performs poorly.

Table S1: Assessment of bias and power approximations. Sample size $n = 1000$ per arm. Two constant infection hazard rates λ_W are considered: (1) $\lambda_W = 0.001$, yielding a 3% attack rate in Arm 0, and (2) $\lambda_W = 0.01$, yielding a 30% attack rate in Arm 0. Incubation period has mean of 6 days (scale parameter = 1). No ramp-up period is included ($R = 0$). No delayed vaccination in Arm 0 ($b = \infty$). Analysis window of $c = 30$ days starting at either $d = 0,\ 6$ or 12. 25,000 simulations run for each scenario. *Power approximation function returns 0.025 assuming other rejection region negligible; we multiply by two. #Expected number of events in Arm 1 is less than 5.

	Infection hazard rate	d	Apparent VE		Power	
			Simulated	Approximated	Simulated	Approximated
$VE_0 = 0\%$	$\lambda_W = 0.001$	0	-0.038	0	0.047	0.050*
		6	-0.039	0	0.048	0.050*
		12	-0.039	0	0.048	0.050*
	$\lambda_W = 0.01$	0	-0.006	0	0.049	0.050*
		6	-0.006	0	0.049	0.050*
		12	-0.007	0	0.049	0.050*
$VE_0 = 50\%$	$\lambda_W = 0.001$	0	0.376	0.400	0.380	0.399
		6	0.464	0.484	0.545	0.562
		12	0.480	0.499	0.578	0.593
	$\lambda_W = 0.01$	0	0.387	0.400	0.990	0.997
		6	0.479	0.484	1.000	1.000
		12	0.496	0.499	1.000	1.000
$VE_0 = 90\%$	$\lambda_W = 0.001$	0	0.708	0.720	0.937	0.928
		6	0.866	0.871	0.974#	0.993
		12	0.895	0.899	0.943#	0.996
	$\lambda_W = 0.01$	0	0.707	0.720	1.000	1.000
		6	0.869	0.871	1.000	1.000
		12	0.898	0.899	1.000	1.000

Table S2: Assessment of bias and power approximations. As in Table S1 but with sample size n=500. *Power approximation function returns 0.025 assuming other rejection region negligible; we multiply by two. #Expected number of events in Arm 1 is less than 5.

	Infection hazard rate	d	Apparent VE		Power	
			Simulated	Approximated	Simulated	Approximated
$VE_0 = 0\%$	$\lambda_W = 0.001$	0	-0.085	0	0.044	0.050*
		6	-0.083	0	0.044	0.050*
		12	-0.086	0	0.043	0.050*
	$\lambda_W = 0.01$	0	-0.011	0	0.050	0.050*
		6	-0.011	0	0.048	0.050*
		12	-0.013	0	0.049	0.050*
$VE_0 = 50\%$	$\lambda_W = 0.001$	0	0.348	0.400	0.196	0.225
		6	0.442	0.484	0.287	0.322
		12	0.456	0.499	0.309	0.341
	$\lambda_W = 0.01$	0	0.384	0.400	0.850	0.916
		6	0.476	0.484	0.964	0.983
		12	0.493	0.499	0.970	0.988
$VE_0 = 90\%$	$\lambda_W = 0.001$	0	0.694	0.720	0.641#	0.678
		6	0.861	0.871	0.735#	0.875
		12	0.891	0.899	0.689#	0.902
	$\lambda_W = 0.01$	0	0.706	0.720	1.000	1.000
		6	0.868	0.871	1.000	1.000
		12	0.898	0.899	1.000	1.000

Table S3: Assessment of bias and power approximations. As in Table S1 but with sample size n=5000. *Power approximation function returns 0.025 assuming other rejection region negligible; we multiply by two.

	Infection hazard rate	d	Apparent VE		Power	
			Simulated	Approximated	Simulated	Approximated
$VE_0 = 0\%$	$\lambda_W = 0.001$	0	-0.005	0	0.049	0.050*
		6	-0.006	0	0.049	0.050*
		12	-0.007	0	0.049	0.050*
	$\lambda_W = 0.01$	0	-0.005	0	0.050	0.050*
		6	-0.005	0	0.052	0.050*
		12	-0.007	0	0.051	0.050*
$VE_0 = 50\%$	$\lambda_W = 0.001$	0	0.396	0.400	0.990	0.968
		6	0.481	0.484	1.000	0.997
		12	0.496	0.499	1.000	0.998
	$\lambda_W = 0.01$	0	0.390	0.400	1.000	1.000
		6	0.482	0.484	1.000	1.000
		12	0.499	0.499	1.000	1.000
$VE_0 = 90\%$	$\lambda_W = 0.001$	0	0.717	0.720	1.000	1.000
		6	0.870	0.871	1.000	1.000
		12	0.898	0.899	1.000	1.000
	$\lambda_W = 0.01$	0	0.709	0.720	1.000	1.000
		6	0.870	0.871	1.000	1.000
		12	0.899	0.899	1.000	1.000

Table S4: Assessment of bias and power approximations. Sample size $n = 1000$ per arm. Two constant infection hazard rates λ_W are considered: (1) $\lambda_W = 0.001$, yielding a 3% attack rate in Arm 0, and (2) $\lambda_W = 0.01$, yielding a 30% attack rate in Arm 0. Incubation period has mean of 6 days (scale parameter = 1). No ramp-up period is included ($R = 0$). Delayed vaccination in Arm 0 after 21 days ($b = 21$). Analysis window of $c = 21$ days starting at either $d = 0,\ 6$ or 12. 25,000 simulations run for each scenario. *Power approximation function returns 0.025 assuming other rejection region negligible; we multiply by two. #Expected number of events in Arm 1 is less than 5.

	Infection		Apparent VE		Power	
	hazard rate	d	Simulated	Approximated	Simulated	Approximated
$VE_0 = 0\%$	$\lambda_W = 0.001$	0	-0.056	0	0.044	0.050*
		6	-0.056	0	0.043	0.050*
		12	-0.057	0	0.044	0.050*
	$\lambda_W = 0.01$	0	-0.009	0	0.048	0.050*
		6	-0.009	0	0.049	0.050*
		12	-0.008	0	0.050	0.050*
$VE_0 = 50\%$	$\lambda_W = 0.001$	0	0.319	0.357	0.219	0.243
		6	0.432	0.465	0.358	0.389
		12	0.376	0.415	0.253	0.282
	$\lambda_W = 0.01$	0	0.343	0.357	0.890	0.953
		6	0.459	0.465	0.990	0.998
		12	0.414	0.415	0.939	0.977
$VE_0 = 90\%$	$\lambda_W = 0.001$	0	0.624	0.643	0.698	0.705
		6	0.845	0.853	0.902#	0.941
		12	0.853	0.863	0.771#	0.885
	$\lambda_W = 0.01$	0	0.629	0.643	1.000	1.000
		6	0.850	0.853	1.000	1.000
		12	0.863	0.863	1.000	1.000

Trials with an unvaccinated control arm

In Figure S10, we assess the apparent VE and apparent power for a trial considering a range of analysis period starts, d, with fixed analysis width, $c = 21$. This example could occur if you only have resources to conduct active follow-up for a fixed period, c. This analysis highlights the impact of the start period d.

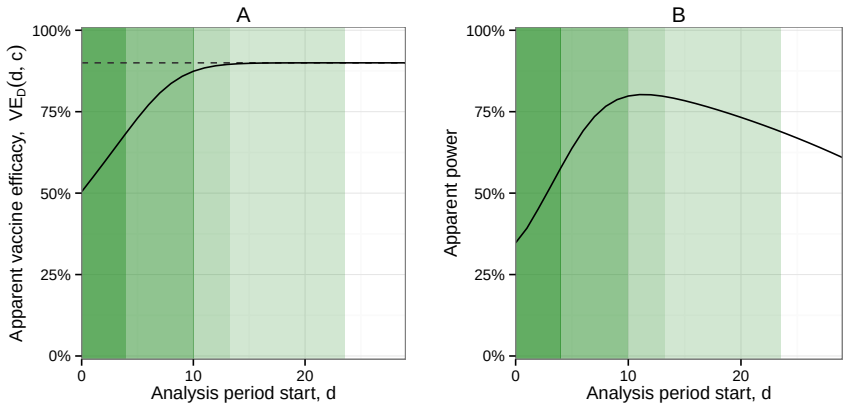

Figure S11: Arm 1 (intervention arm) vaccinated on day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Background infection hazard rate decreasing linearly from $\lambda_W(0) = 0.0015$ to $\lambda_W(60) = 0$. Standard setting. Sample size $n = 500$ per arm. (Setting as in Figure S8.) Analysis period $[d, d + 21)$ with width $c = 21$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

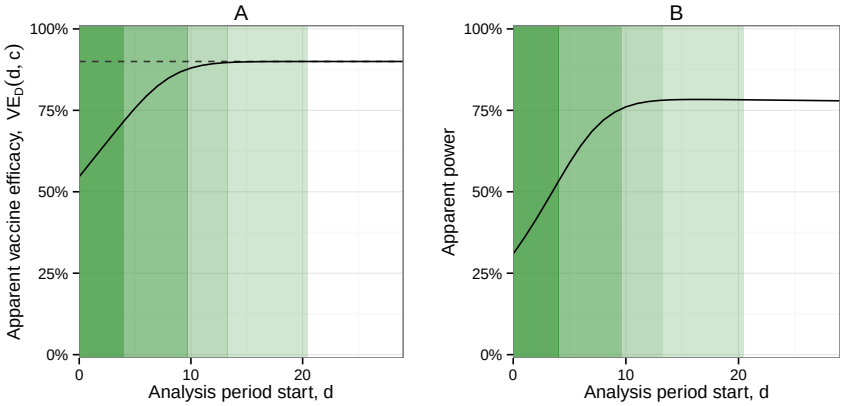

Figure S10: Arm 1 (intervention arm) vaccinated own day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Constant background infection hazard rate $\lambda_W = 0.001$. Standard setting. Sample size $n = 500$ per arm. (Setting as in Figure 2.) Analysis period $[d, d + c)$ with width $c = 21$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

In Figure S11, we show results for apparent VE and power in the setting of a decreasing background infection hazard rate. A drop in infection hazard rate could occur because of disease prevention interventions, increased awareness leading to modified behavior, and exhaustion of susceptibles. The results for apparent VE are identical to those in Figure S10, but when the background infection hazard rate is decreasing, starting the analysis period too late can result in a loss of power because of a lack of events.

In Figure S12, we consider here the impact decreasing background infection hazard rate when a a fixed analysis period end ($d + c = 50$) is used. We can directly compare with Figure 5 which uses the same setting

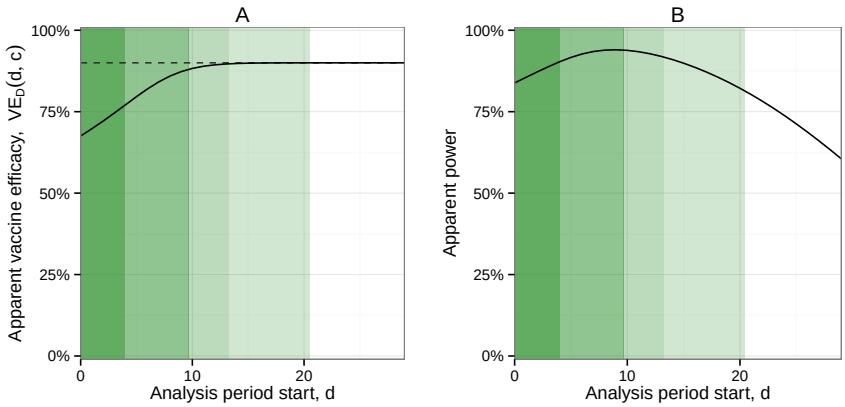

Figure S12: Arm 1 (intervention arm) vaccinated on day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Background infection hazard rate decreasing linearly from $\lambda_W(0) = 0.0015$ to $\lambda_W(60) = 0$. Standard setting. Sample size $n = 500$ per arm. (Setting as in Figure S8.) Analysis period $[d, 50)$ with width $c = 50 - d$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

except with a constant background infection hazard rate. We note slightly more bias for the earliest analysis periods. With a higher infection hazard rate early in the trial, more events are observed in the early period before the immune response has developed. In addition, a more dramatic drop in power is observed for later d values because of the decline in events.

In Figure S13, we consider an incubation period with the same mean but a wider distribution. This figure can be best compared to Figure S10. We note a longer period before VE and power stabilize. In addition, at the earliest values of d, we see greater bias and lower power.

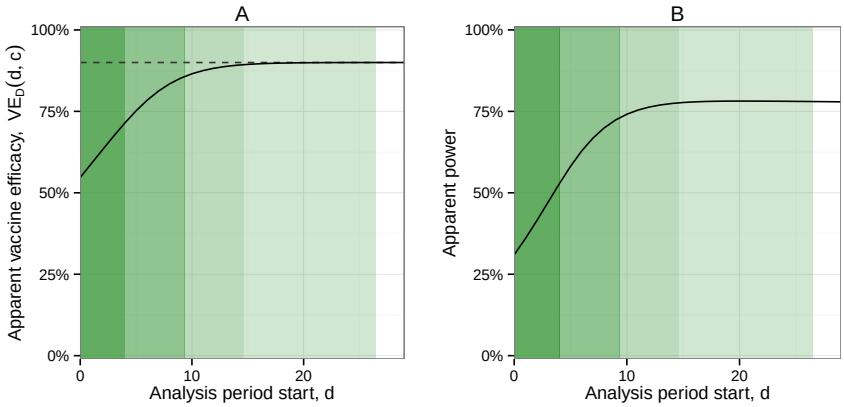

Figure S13: Arm 1 (intervention arm) vaccinated own day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Constant background infection hazard rate $\lambda_W = 0.001$. Standard setting except more diffuse incubation period distribution (same mean = 6, but scale = 2). Sample size $n = 500$ per arm. (Setting as in Figure 2.) Analysis period $[d, d + c)$ with width $c = 21$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

In Figure S14, we consider an incubation period with a greater mean (12 as compared to 6 days). This figure can be best compared to Figure S10. We note a longer period before VE and power stabilize. In addition, at the earliest values of d, we see greater bias and lower power.

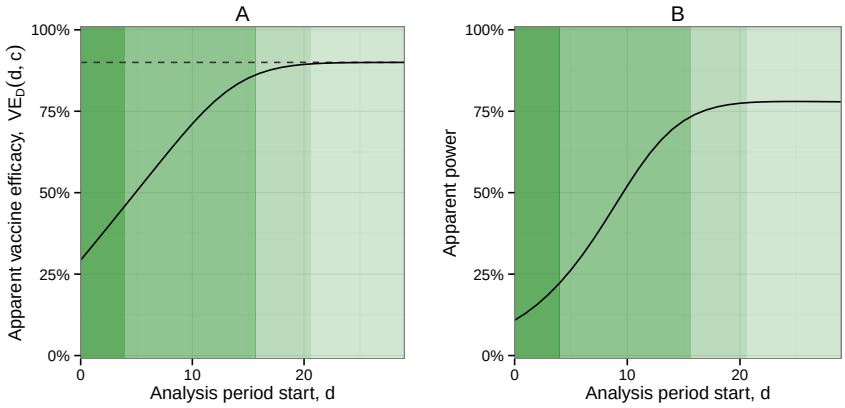

Figure S14: Arm 1 (intervention arm) vaccinated own day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Constant background infection hazard rate $\lambda_W = 0.001$. Standard setting except more diffuse incubation period distribution (mean = 12, scale = 1). Sample size $n = 500$ per arm. (Setting as in Figure 2.) Analysis period $[d, d + c)$ with width $c = 21$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

In the three figures below, we vary the ramp-up period length R. Figure S15 has ramp-up period $R = 0$. Figure S16 has ramp-up period $R = 6$. Figure S17 has ramp-up period $R = 12$. The ramp-up period shifts the optimal analysis period forward, though as the ramp-up period increases, the period of partial efficacy lengthens. Thus, especially for maximizing power, the optimal d might be earlier than would be predicted by the increase in R.

In Figure S18, we present three different shapes for the ramp-up period (fast growth, linear growth, and slow growth). They perform very similarly, but fast growth is associated with slightly lower VE and slightly higher power.

In Figure S19, we consider the impact of the underlying vaccine efficacy VE_0. The optimal value d is similar across all levels of VE_0, though it is slightly later when VE_0 is highest. This occurs because high VE_0 requires the largest change in illness onset hazard ratio, meaning that the period of partial efficacy can induce the most bias in apparent VE.

In Figure S20 we consider the impact of variable background infection hazard rate λ_W. While the rate does not impact bias, it has a dramatic impact on the absolute level power as expected. Notably, for the highest

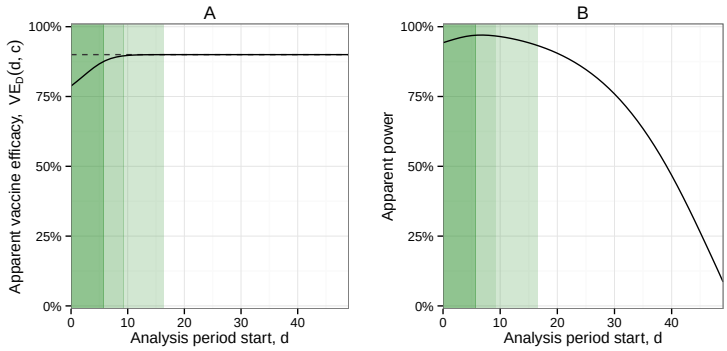

Figure S15: Arm 1 (intervention arm) vaccinated on day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Constant background infection hazard rate 0.001. Standard setting, except $R = 0$. Sample size $n = 500$ per arm. (Setting as in Figure S8.) Analysis period $[d, 50)$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

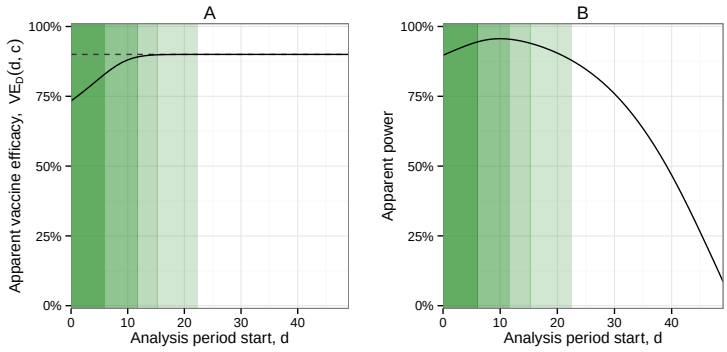

Figure S16: As Figure S15 with $R = 6$

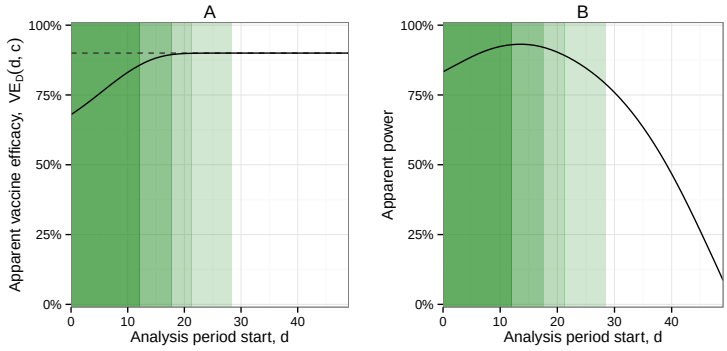

Figure S17: As Figure S15 with $R = 12$

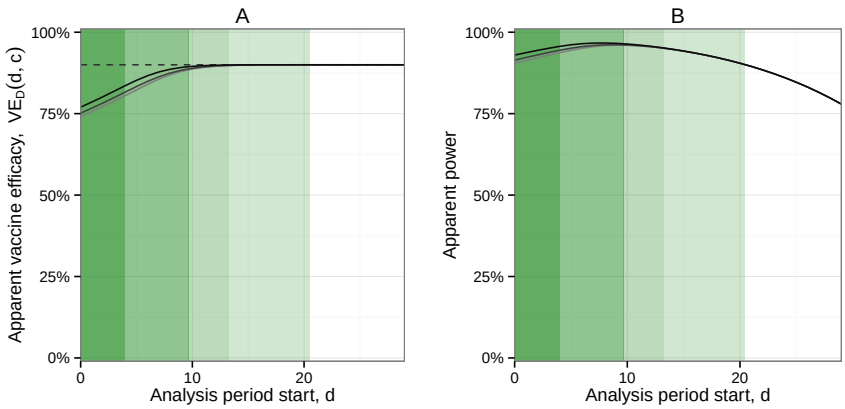

Figure S18: Arm 1 (intervention arm) vaccinated on day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Constant background infection hazard rate 0.001. Standard setting, except for variable ramp-up period shape. From darkest to lightest: fast growth, linear growth, slow growth. Sample size $n = 500$ per arm. (Setting as in Figure S8.) Analysis period $[d, d + 21)$ with width $c = 21$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

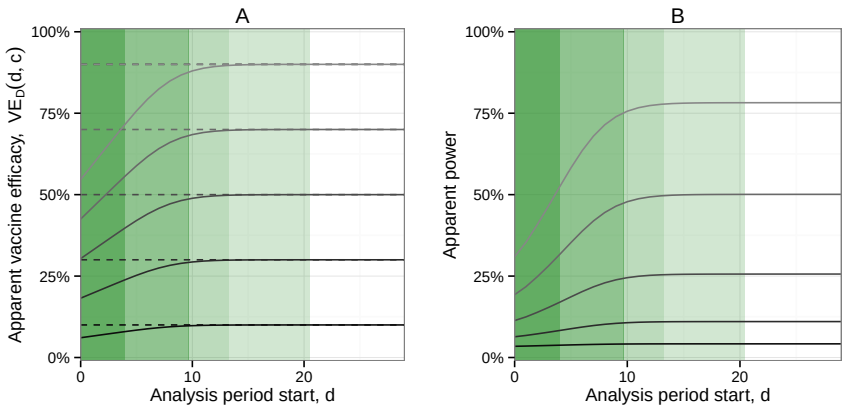

Figure S19: Arm 1 (intervention arm) vaccinated on day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Constant background infection hazard rate 0.001. Includes VE 10, 30, 50, 70, 90%. Standard setting. Sample size $n = 500$ per arm. (Setting as in Figure S8.) Analysis period $[d, d + 21)$ with width $c = 21$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

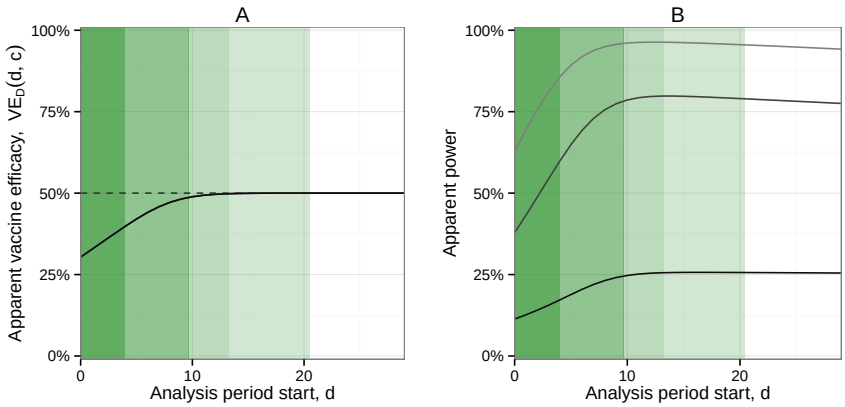

Figure S20: Arm 1 (intervention arm) vaccinated on day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Constant background infection hazard rate of 0.001, 0.005, and 0.01 considered. VE=50%. Standard setting. Sample size $n = 500$ per arm. (Setting as in Figure S8.) Analysis period $[d, d + 21)$ with width $c = 21$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

rate, we can see a slight decrease in power over time due to exhaustion of susceptibles in the population who must be censored, decreasing the overall sample size.

Trials with a delayed vaccination arm

In Figure S21, a delayed vaccination design with a declining background infection hazard rate is considered. The scenario can be directly compared to Figure6 , which has a constant background infection hazard rate.

In Figure S22, the same setting as Figure 6 is presented but with a sample size of $n = 1000$ per arm. As expected, the power is higher.

In Figure S23, we consider the setting where a stringent cutoff is applied to end the analysis period at the time of delayed vaccination. In this scenario, the apparent vaccine efficacy does not decrease for large d, but the power is so low that this approach has limited value in this scenario. It may perform better in other settings, especially those with a longer vaccination delay.

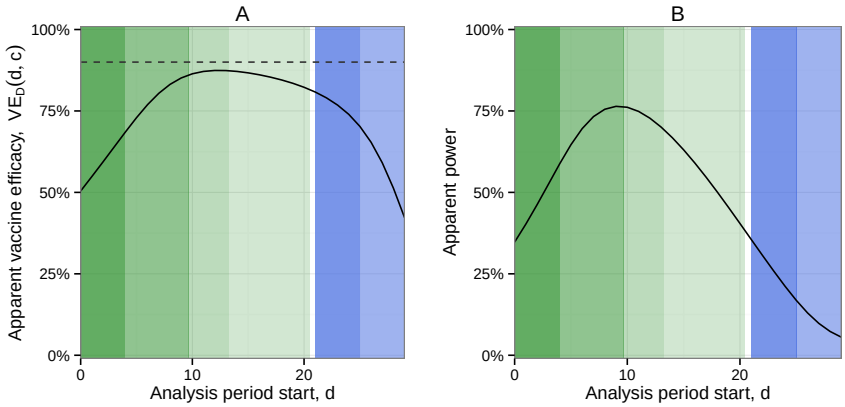

Figure S21: Arm 1 (immediate vaccination arm) vaccinated on day $s_1 = 0$; Arm 0 (delayed vaccination arm) vaccinated on day $s_1 = 21$ ($b = 21$). Background infection hazard rate decreasing linearly from $\lambda_W(0) = 0.0015$ to $\lambda_W(60) = 0$. Standard setting. Sample size $n = 500$ per arm. (Setting as in Figure S9.) Analysis period $[d, d+21)$ with width $c = 21$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

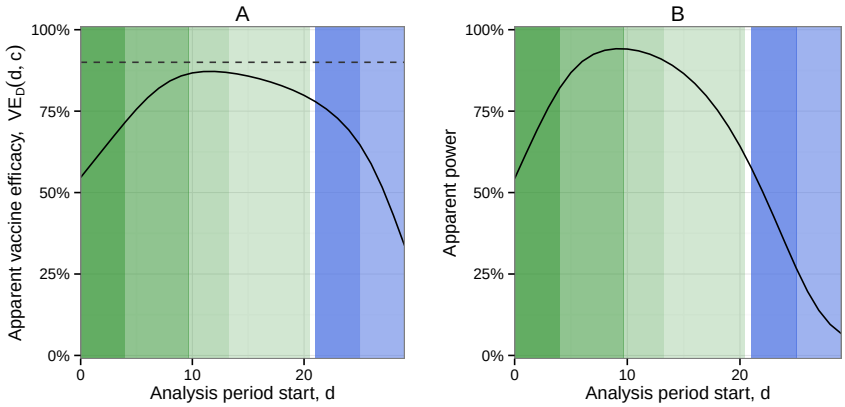

Figure S22: As in Figure S24, but with sample size $n = 1000$ per arm. (A) Apparent VE to assess bias. (B) Apparent power.

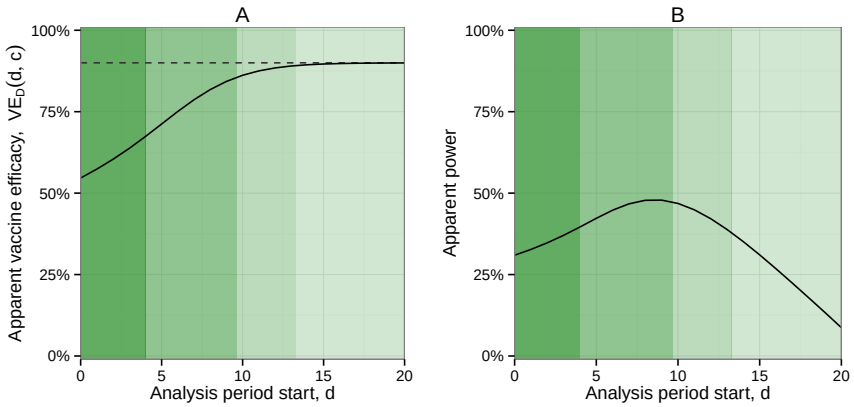

Figure S23: Arm 1 (immediate vaccination arm) vaccinated on day $s_1 = 0$; Arm 0 (delayed vaccination arm) vaccinated on day $s_1 = 21$ ($b = 21$). Constant background infection hazard rate $\lambda_W = 0.001$. Standard setting. Sample size $n = 500$ per arm. (Setting as in Figure S9.) Analysis period ends at time of delayed vaccination, $[d, b)$ with width $c = b - d$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

To improve the low power observed in Figure S23, one option would be to extend the analysis end period to be the time of delayed vaccination plus the mean or median incubation period. Note that this approach may include some vaccine protected time in the delayed arm if the vaccine is very fast-acting and/or there is a post-exposure prophylactic effect. In general, though, infections during this period likely occurred while the delayed arm was unvaccinated. In Figure S24, an analysis period end on day 27 (21+6) is considered, where $b = 21$ and the mean incubation period is 6 days. We see very little change in the bias but a notable improvement in the power.

In Figure S25, we consider the impact of the underlying vaccine efficacy VE_0. The optimal value d is similar across all levels of VE_0, though it is slightly later when VE_0 is highest. This occurs because high VE_0 requires the largest change in illness onset hazard ratio, meaning that the period of partial efficacy can induce the most bias in apparent VE. It is also noteworthy that when $VE_0 = 100\%$, larger values of d do not induce bias in apparent VE as no additional cases are observed in the immediately vaccinated arm.

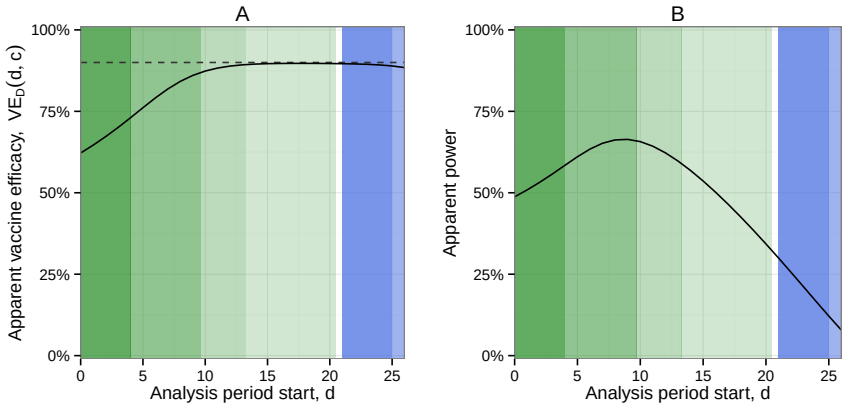

Figure S24: Arm 1 (immediate vaccination arm) vaccinated on day $s_1 = 0$; Arm 0 (delayed vaccination arm) vaccinated on day $s_1 = 21$ ($b = 21$). Constant background infection hazard rate $\lambda_W = 0.001$. Standard setting. Sample size $n = 500$ per arm. (Setting as in Figure 3.) Analysis period ends at time of delayed vaccination, $[d, b + 6)$ with width $c = b - d - 6$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.

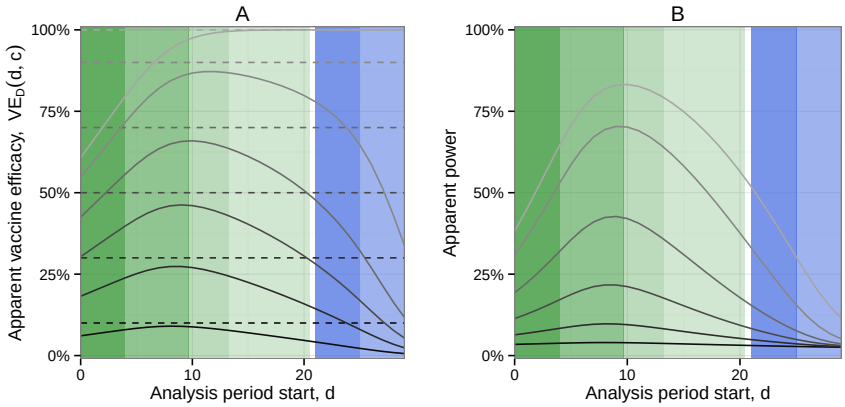

Figure S25: Arm 1 (intervention arm) vaccinated on day $s_1 = 0$; Arm 0 (comparator arm) never vaccinated ($b = \infty$). Constant background infection hazard rate 0.001. From darkest to lightest, includes $VE_0 = 10\%$, 30%, 50%, 70%, 90%, and 100%. Standard setting. Sample size $n = 500$ per arm. Analysis period $[d, d+21)$ with width $c = 21$; a range of d values are considered. (A) Apparent VE to assess bias. (B) Apparent power.