Narrative Review
Acta Medica Academica 2022;51(2):134-146
DOI: 10.5644/ama2006-124.383

On Measuring Vaccine Effectiveness with Observational Study Designs

Farrokh Habibzadeh[1,*], Parham Habibzadeh[2], Mahboobeh Yadollahie[3]

[1]Global Virus Network, Middle East Region, Shiraz, Iran, [2]Division of Clinical Care and Research, Institute of Human Virology, University of Maryland School of Medicine, Baltimore, USA, [3]Freelance Researcher, Shiraz, Iran

Correspondence: *Farrokh.Habibzadeh@gmail.com*; Tel.: + 98 71 32252258

Received: 2 July 2022; **Accepted:** 31 August 2022

Abstract
Herein, we present a bird's eye view of common observational study designs utilized for measurement of vaccine effectiveness. Assessing vaccines effectiveness is an integral part of vaccine research, particularly for the newly developed vaccines. A cohort study is prospective, directing from an exposure to one or more outcomes. The design is the best method to ascertain the attack rate of an infectious disease. A traditional case-control study is retrospective, directing from a given outcome to one or more exposures. The design cannot provide the relative risk, but it can provide the odds ratio, which is a good estimation of the relative risk when the attack rate is low. Critically depending on laboratory test results and performance, the test-negative case-control study design is another type of observational study commonly used nowadays for the evaluation of the vaccine effectiveness. Comparing to cohort and traditional case-control designs, conducting a test-negative case-control study is relatively cheaper and faster. Herein, we describe each of the above-mentioned study designs through examples generated by a Monte-Carlo simulation program assuming real-world conditions. **Conclusion.** The simulation shows that regardless of the study design employed, the diagnostic test specificity is of utmost importance in providing a valid estimate of the vaccine effectiveness.

Key Words: Vaccines ▪ Research Design ▪ Cohort Studies ▪ Case-Control Studies ▪ SARS-CoV-2.

Introduction

For every new vaccine, the vaccine effectiveness (*VE*), an index reflecting the measure of infection or disease risk reduction attributable to vaccination among vaccinated individuals compared with unvaccinated people under real-world situations, should be determined to figure out the future vaccination policy and strategy to be implemented. The *VE* is defined as:

$$VE = \frac{AR_{unvac} - AR_{vac}}{AR_{unvac}} = 1 - \frac{AR_{vac}}{AR_{unvac}} = 1 - RR \quad \text{(Eq 1)}$$

where AR_{unvac} and AR_{vac} represent the attack rates of the infection in the unvaccinated and vaccinated individuals, respectively, and *RR* is the relative risk (1, 2). The gold standard study designs to determine the *VE* are the randomized clinical trial and cohort studies (3). In such studies, the *AR* of the infection is evaluated after a period of time, say 3 months after vaccination, in the two groups of vaccinated and unvaccinated individuals who had been either randomized into two groups (in a randomized clinical trial) to abolish the effect of confounding variables, or at least matched for known important confounders (in a cohort study).

Although clinical trials are the best study design to measure *VE*, under certain circumstances, for ethical concerns, we are not able to conduct clinical trials; for instance, it is unethical not to vaccinate a susceptible person with exposure to

* ORCID: 0000-0001-5360-2900

an infectious agent for which an approved vaccine is available. Therefore, we need to conduct observational studies to determine the *VE* under more realistic situations in wider population groups in whom the *VE* may differ from that observed in clinical trials for several reasons including difference in the geographic parameters, sub-populations not included or under-represented in the original trials (*e.g.*, children and pregnant women), sub-optimal consideration of the cold-chain necessary for optimal effectiveness of the vaccine, incomplete vaccination, incorrect time spacing between the vaccine doses administered, emergence of new variants of the infectious agent, and many other factors (4, 5).

In this review, we will discuss the *pros* and *cons* of common observational studies used to determine *VE*. To better understand the issue, we first overview the study designs, present examples through using a simulation program, and finally describe a new type of case-control study, the so-called test-negative case-control (TNCC) study that has become common to use for the assessment of *VE*.

Setting

Suppose we want to study an arbitrarily chosen population of 2 000 000 individuals (for example, the population of a city like Shiraz) and that we have vaccinated a hypothetical fraction of 40% of the people against an infectious agent. Let the true *VE* after 3 months of vaccination be 0.70 and that we use a diagnostic test with a test sensitivity (*Se*) of 60% and a specificity (*Sp*) of 100% for the detection of the infection. Furthermore, suppose that the infected people present with signs and symptoms similar to a flu and that there is another flu-like illness that might also affect people living in the study community, independent of whether they have already been affected by the infection of interest or not. Also, assume that the infection has an *AR* of 15% in unvaccinated individuals (consistent with the *AR* of influenza) (6); and that the *AR* of the flu-like illness is 30% (consistent with the attack rates of non-influenza flu-like illness seen during a cold season) (7, 8). To make things simple, let us assume that the *AR* of the flu-like illness is independent of vaccination status against the infection of interest, duration since vaccination, age of people, and other variables. Now, suppose we want to estimate the *VE* of an influenza vaccine using various observational study designs.

Observational Studies

Clinical research studies can be classified into two broad categories — observational and interventional studies (9). In observational studies, the researcher just observes; no intervention occurs. Observational studies can further be categorized into longitudinal and cross-sectional studies, based on how the observations are made over the study period. Two important longitudinal observational studies are cohort and case-control studies (9).

Cohort

In a cohort study, two groups of individuals with and without exposure to a certain agent (*e.g.*, a vaccine) are followed for a certain period of time. The two groups are similar in (theoretically, perfectly matched for) all other variables but their exposure (10). In its simplest form, we compare the experience of the exposed group with that of the unexposed group and measure the incidence of a certain outcome in the two groups. If the incidence of the outcome in one group is significantly different from that in other group, then we conclude that there should be an association between the exposure and outcome of interest (10). Cohort studies are prospective — directing from an exposure to one or more outcomes.

Suppose that the disease of interest is influenza, and that a random sample of 50 vaccinated (exposed) and 50 unvaccinated (unexposed) individuals was taken from the above-mentioned population (Figure 1A). For the sake of simplicity, let us assume that the two groups were perfectly matched for other variables. The first five columns (Figure 1A) represent the exposed individuals; the remaining, unexposed. Let us define the outcome of interest as presence of influenza ascertained by

a diagnostic test (a positive test). Suppose that the 100 study participants were followed for 3 months and that 10 of whom were found test positive, which translated into an *AR* of 10% in 3 months. In other words, the risk of infection within 3 months, regardless of vaccination status, was 10%. Two (4%) out of 50 of the exposed (vaccinated) participants and 8 (16%) of 50 of unexposed (unvaccinated) individuals developed the outcome of interest (a positive test result). The *AR* of the infection was therefore 4% in the exposed and 16% in the unexposed group (Table 1B). The unexposed group carried a 4-fold (= 16%/4%) increase in the risk of infection as compared with the vaccinated group. In other words, vaccination decreased the risk of infection by 75%, the estimated *VE*.

Table 1. Test Results Stratified by Vaccination Status in Various Study Designs

A) General template	Disease*		
	Present	Absent	
Vaccinated	*a*	*b*	*a+b*
Unvaccinated	*c*	*d*	*c+d*
	a+c	*b+d*	*n*

B) Cohort	Disease*		
	Present	Absent	
Vaccinated	2	48	50
Unvaccinated	8	d	50
	10	90	100

C) Traditional Case-Control	Disease*		
	Present	Absent	
Vaccinated	8	23	
Unvaccinated	42	27	
	50	50	100

D) Test-Negative Case-Control	Test		
	Positive	Negative	
Vaccinated	2	32	
Unvaccinated	11	55	
	13	87	100

A) a 2×2 contingency table, the general template for various study designs of *n* participants, and examples for samples of 100 individuals using B) cohort; C) traditional case-control; D) test-negative case-control studies associated with Figure 1; *The status was determined by using a diagnostic test (or a battery of tests).

Let us examine the general parametric form of a cohort study of *n* participants (Table 1A). Then we can write:

$$AR = \frac{\text{number of infected people}}{\text{total number of participants at risk of infection}} = \frac{a+c}{n} = \frac{a+c}{a+b+c+d} \quad \text{(Eq 2)}$$

where *AR* is the marginal risk of infection in the whole study participants, regardless of vaccination status, and *a*, *b*, *c*, and *d* are Table cell values (Table 1A). In the same way, we can calculate the *AR* in the vaccinated and unvaccinated groups:

$$AR_{vac} = \frac{\text{number of vaccinated people with infection}}{\text{total number of vaccinated individuals at risk of infection}} = \frac{a}{a+b} \quad \text{(Eq 3)}$$

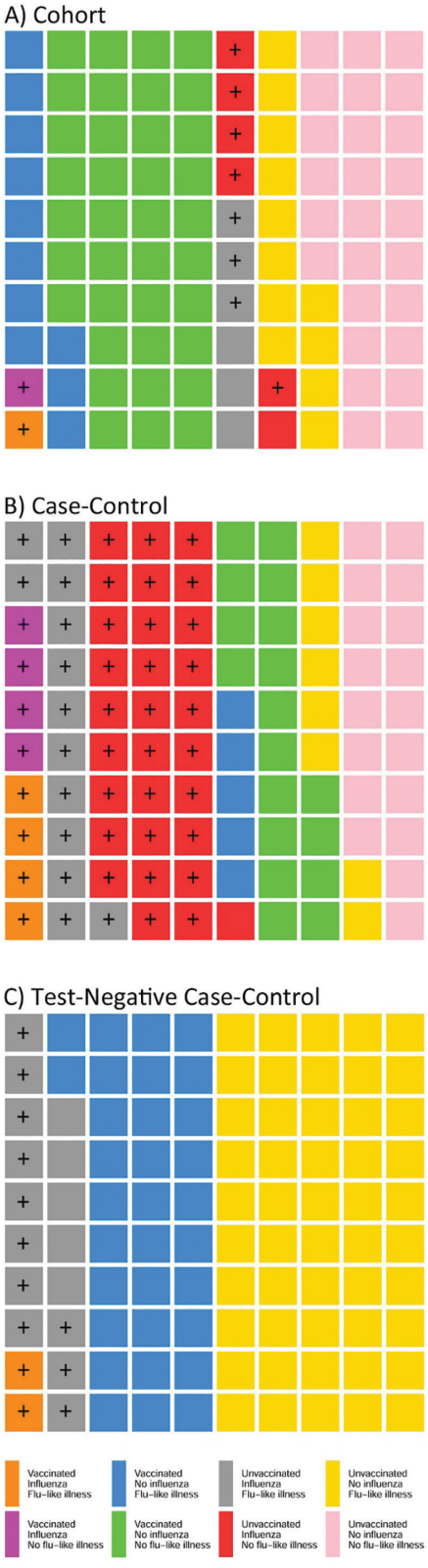
and

$$AR_{unvac} = \frac{\text{number of unvaccinated people with infection}}{\text{total number of unvaccinated individuals at risk of infection}} = \frac{c}{c+d} \quad \text{(Eq 4)}$$

The *RR* is:

$$RR = \frac{\text{risk in the exposed group}}{\text{risk in the unexposed group}} = \frac{AR_{vac}}{AR_{unvac}} = \frac{a/(a+b)}{c/(c+d)} = \frac{a(c+d)}{c(a+b)} \quad \text{(Eq 5)}$$

When the outcome is something like hospitalization or death, recognizing the outcome is easy. However, when the outcome is a variable like presence or absence of a certain disease, identifying the outcome is not always that easy. Normally, we rely on diagnostic test (or a battery of tests) results



to classify people into diseased and undiseased groups. Depending on the distribution of test results, test *Se* and *Sp*, and the prevalence of the disease, levels of uncertainty would be introduced. In our example, we missed 4 cases of influenza (3 gray and 1 red squares in Figure 1A) for having false-negative test results. This would affect the estimated risks in the two study groups and the *VE*. Having a test *Sp* of 100%, no false-positive result occurred.

The required minimum sample size for a cohort study being conducted for the measurement of *VE* is a function of the anticipated *VE* and the desired width of its confidence interval, the percentage of the vaccine coverage, and presumed *AR* in the unvaccinated group (5, 11). The minimum sample size can be calculated using a calculator available online (12).

A cohort study has several advantages. It is the best type of study design for the measurement of the incidence (*AR*, incidence, risk) of a given outcome (*e.g.*, a disease) under certain conditions. The temporal sequence of events is typically clear and one can usually make a cause-and-effect inference and ascertain the natural history of a disease. With this design, it is possible to measure risks of several outcomes (*e.g.*, infection, hospitalization, and death) and their association with a given exposure (*e.g.*, vaccination). The design is also very good to assess the risk of rare exposures (10). However, it

Figure 1. Examples generated by the simulation program based on the assumptions made in the Setting section of the article: 100-individual samples taken from a 2 000 000-individual population with a vaccination coverage of 40%. The test had a sensitivity of 60% and specificity of 100% for the detection of influenza (the outcome); attack rate of 15% in unvaccinated individuals for influenza, (6) and 30% for the flu-like illness (7, 8). A) Cohort study design: The left-most 5 columns are vaccinated; the remaining, unvaccinated. Those with a positive test are indicated. Others were test-negative. 4 individuals with influenza (3 gray and 1 red squares) had false-negative test results. B) Traditional case-control study design: The left-most 5 columns are test-positive individuals (considered cases); others were test-negative (controls). One of the test-negative individuals in the control group (1 red square) had really influenza (a false-negative test result). C) Test-negative case-control study design: Note that all 100 individuals have flu-like illness. 5 individuals with influenza (5 gray squares) had false-negative test results. The test specificity was 100%, hence, no false-positive result was obtained.

has some disadvantages. Following a large cohort for a long period of time is expensive. The intensity of follow-up should be equal in the two study groups, which is not always possible. Many of the study participants may be lost to follow up. Some of the participants in one group may decide to switch to another group (*e.g.*, an unvaccinated person may decide to receive the vaccine). Matching of the study groups for the important covariates and controlling of the known confounding variables are not always easy. And, the design is not appropriate for rare outcomes (10). Generally, cohort studies are more efficient in situations where the incidence of outcome is higher than the prevalence of exposure (13).

Case-Control

Case-control study design is retrospective. It begins with a certain outcome (*e.g.*, presence of an infection) and returns back in time to examine the level of exposure to one or more factors (*e.g.*, vaccination status) of those with the outcome (cases) and without the outcome (controls). The cases and controls are similar in (theoretically, perfectly matched for) all other variables but the status of the outcome (13). In a case-control study, because we are typically not aware of the real proportion of the cases and controls in the population, we are not able to estimate the risk (incidence and *AR*) in the two groups and the *RR*; instead, we compare odds of exposure in the two groups and calculate the odds ratio (*OR*), as follows (Table 1A):

$$\begin{aligned} OR &= \frac{\text{odds of exposure in cases}}{\text{odds of exposure in controls}} \\ &= \frac{a/c}{b/d} = \frac{a/b}{c/d} \\ &\approx \frac{a/(a+b)}{c/(c+d)}, \text{ if } a \ll b \wedge c \ll d \\ &\approx \frac{a(c+d)}{c(a+b)} \\ &\approx RR \end{aligned} \qquad \text{(Eq 6)}$$

If the *AR* is low (*i.e.*, $a \ll b \wedge c \ll d$), then $b \approx a+b$ and $d \approx c+d$, *OR* is an acceptable estimation for *RR* (Eq 6), and *VE* can be calculated as 1 – *OR* (2).

Suppose a random sample of 50 diseased (defined as test-positive individuals) and 50 undiseased (test-negative) people from the above-mentioned population was taken (Figure 1B, Table 1C). The *OR* (Eq 6) is then:

$$\begin{aligned} OR &= \frac{a/b}{c/d} \\ &= \frac{8/23}{42/27} = \frac{0.35}{1.56} = 0.22 \end{aligned} \qquad \text{(Eq 7)}$$

The estimated *VE* after 3 months of vaccination was then 0.78 (= 1 – 0.22). Depending on the test results distribution and performance, and the prevalence of the disease of interest, false-positive and false-negative results may occur. In our example, we had one false-negative individual (Figure 1B).

The required minimum sample size for a case-control study being conducted for the measurement of *VE* is a function of the anticipated *VE* and the desired width of its confidence interval, and the presumed prevalence of vaccination in the control (undiseased) group (5, 11). The minimum sample size can be calculated using a calculator available online (12).

The case-control design has several advantages. It is the most efficient study design in terms of time and money spent, and efforts made (13). The design is especially appropriate for rare outcomes; this makes sense if we consider that for observing a rare outcome, a researcher conducting a cohort study should have normally follow a large group of people for a long period of time. In a case-control study we can begin with a given outcome and examine one or more exposures. However, the design is not appropriate when the exposure frequency is low. In measuring the *VE*, the *AR* is usually acceptably low and case-control studies would give satisfactory estimates. Although it is usually easier to conduct a case-control study compared to a cohort

study, conducting the investigation could still be labor-intensive — intense efforts are still needed to identify and recruit the controls from the population (13).

Test-Negative Case-Control (TNCC)

A TNCC study design technically has a case-control design. The only difference is the way the cases and controls are recruited. It seems that the first complete description of the method dates back to 1985, when Miettinen mentioned the methodology in his book (14). The method has widely been used for measuring influenza *VE*, first employed by Skowronski in Canada in 2004–2005 (15). The design has been frequently used thereafter, most commonly for assessing the *VE* (16-20).

In a TNCC study, a group of people referred to a healthcare center for a reason, say complaining of flu-like illness, is considered the study sample. All test-positive individuals are considered "cases;" test-negative individuals, "controls." The exposure (vaccination) status is then ascertained in the two groups (21).

Suppose a sample of 100 patients complaining of a flu-like illness referred to a healthcare center was taken (Figure 1C, Table 1D). A diagnostic test was then performed for each of the study participants to ascertain whether they had the infection of interest or not. The 13 test-positive individuals (2 orange and 11 gray squares in Figure 1C) were considered cases; the remaining 87 (= 100 – 13), controls. The individuals were then asked about their vaccination status (exposure, Table 1D). As in traditional case-control study design, the *OR* can be calculated as:

$$OR = \frac{a/b}{c/d} = \frac{2/32}{11/55} = \frac{0.06}{0.20} = 0.31 \quad \text{(Eq 8)}$$

The *VE* after 3 months of vaccination was then 0.69 (= 1 – 0.31).

The TNCC design has all the advantages and disadvantages of traditional case-controls (*e.g.*, outcome misclassification, recall bias). We observed 5 diseased individuals with false-negative test results, misclassified to the control group (Figure 1C). The study design, however, has the advantage of reducing the effect of difference in health-care seeking behavior between the exposed and unexposed groups (3-5, 16). TNCC study is relatively cheaper and faster to conduct in comparison with cohort and traditional case-control studies (16).

Monte-Carlo Simulation

R software version 4.1.0 (R Project for Statistical Computing) was used for simulation. The pseudocode of the program is presented in Table 2 (see Appendix for the R codes). The simulation parameters were initially set to the values described earlier (see Setting).

The estimated *VE* derived from a study with a sample size of 100, although more appropriate for presentation as a graph (Figure 1), would be associated with high degrees of uncertainties due to sampling error. The simulation program was therefore run with a sample size of 10 000 individuals, keeping other parameters the same (Table 2), which resulted in estimated *VE* values of 0.686, 0.705, and 0.710 for cohort, traditional case-control, and TNCC designs, respectively (Table 3).

To obtain more accurate results, we repeated the above *in silico* experiments of 10 000 individuals for different combinations of the test *Se* (varying from 60% to 100%) and *Sp* (varying from 80% to 100%) for 1000 times, and reported the mean of *VE* (Figure 2). The simulation was performed for *AR* of 15% for unvaccinated individuals, compatible with the *AR* of influenza (6); and an *AR* of 5% for SARS-CoV-2 infection (5).

As expected, cohort design gave the most accurate estimates provided using a highly specific diagnostic test. Results obtained from traditional case-control design and TNCC design were very similar (Figure 2). The test *Sp* was more important than the *Se*.

Table 2. The Pseudocode of the Simulation Program*

```
Begin
  Make a population of 2 000 000 individuals, 40% of whom were vaccinated;
  15%, had the disease; and 30%, had a flu-like illness
  Loop for various values of Se and Sp
    Determine the test status of each individual in the population based on Se and Sp
    # Cohort
    Loop for 1000 times
      Choose at random from the population, 5000 vaccinated and 5000 unvaccinated individuals
      Calculate the RR based on the test status
      VE = 1 – RR
    End Loop
    # Traditional Case-Control
    Loop for 1000 times
      Choose at random from the population, 5000 diseased and 5000 undiseaded individuals
      Calculate the OR based on the test status
      VE = 1 – OR
    End Loop
    # Test-Negative Case-Control
    Loop for 1000 times
      Choose at random from the population, 10 000 individuals with flu-like illness
      Calculate the OR based on the test status
      VE = 1 – OR
    End Loop
  End Loop
  Draw the results as a graph
End
```

*See Appendix for the R codes.

Table 3. Test Results Stratified by Vaccination Status in 10 000-individual Samples Using Different Study Designs

A) Cohort	Disease*		
	Present	Absent	
Vaccinated	136	4864	5000
Unvaccinated	433	4567	5000
	569	9431	10 000

B) Traditional Case-Control	Disease*		
	Present	Absent	
Vaccinated	874	2089	
Unvaccinated	4126	2911	
			10 000

C) Test-Negative Case-Control	Test		
	Positive	Negative	
Vaccinated	113	3901	
Unvaccinated	543	5443	
			10 000

A) cohort; B) traditional case-control; C) test-negative case-control; *The status was determined by using a diagnostic test (or a battery of tests).

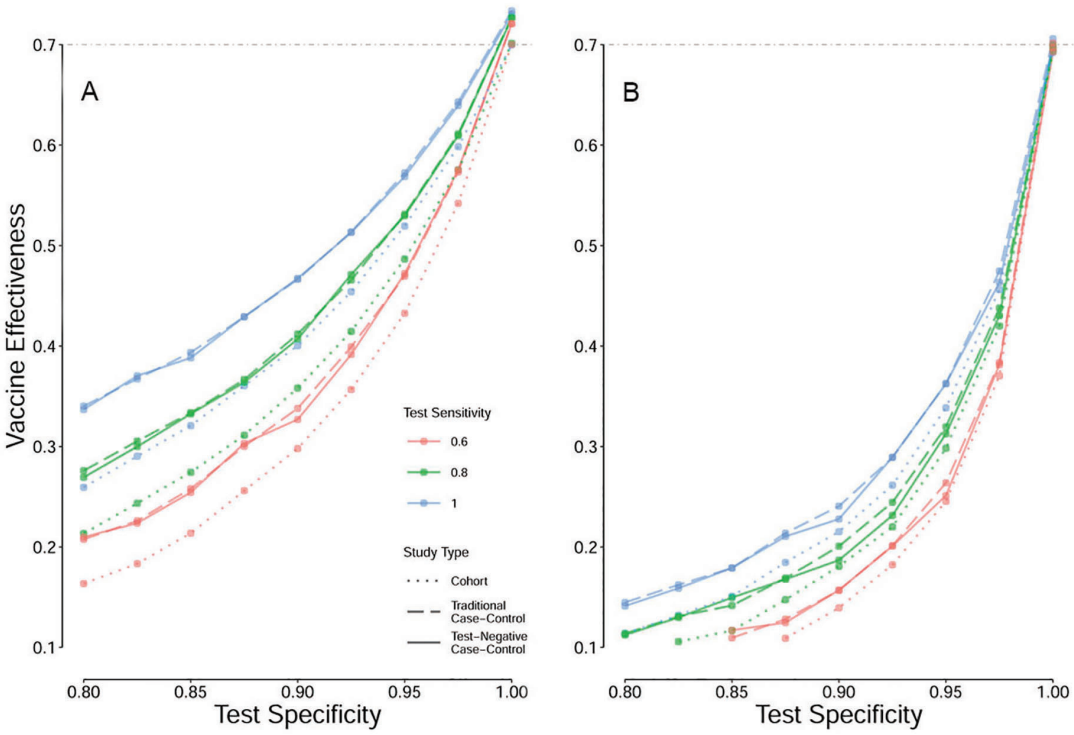


Figure 2. Vaccine effectiveness derived from *in silico* studies simulating various study types under different conditions. The horizontal dash-dotted gray line represents the true vaccine effectiveness. The results are mean values of the vaccine effectiveness derived from 1000 repetitions of a Monte-Carlo simulation (Table 2). Each time, 10 000 individuals were examined under combinations of the test sensitivity (varying from 0.6 to 1 [*i.e.*, 60% to 100%]) and specificity (varying from 0.8 to 1 [*i.e.*, 80% to 100%]). The simulation was performed for attack rates of A) 15% for unvaccinated individuals, compatible with the that of influenza (6); and B) 5%, for SARS-CoV-2 infection (5). In both situations, the results obtained are satisfactory when the test specificity is almost 1 (*i.e.*, 100%).

Discussion

It was found that cohort studies are superior to other observational designs for the evaluation of *VE*. This is because only a cohort study can accurately determine the *AR* of the disease in the vaccinated and unvaccinated groups. A case-control study can only provide *OR* as an estimate for *RR*. However, this assumption is only true when the *AR* is considerably low (Eq 6). Given the low *AR* for many infectious disease (*e.g.*, influenza and SARS-CoV-2), the calculated *OR* is not much different from the *RR* and thus a case-control study can also give a satisfactory estimate for the *VE*. Furthermore, using logistic regression analysis, the *OR* can be adjusted for various confounding variables to give an adjusted *VE* (22). This is in keeping with our results; for an *AR* of 15% (Figure 2, left panel), the *VE* estimates derived from the two studied case-control designs overestimated the true *VE*; with a lower *AR* of 5%, these studies gave estimates closer to the real *VE*.

The TNCC design has the advantage over cohort and traditional case-control studies for being less expensive and faster to conduct (16). In keeping with our simulation results, numerous *in silico* studies have also shown that the design can provide estimates of *VE* in good agreement with those of cohort and traditional case-control studies provided that a highly specific diagnostic test is used (2, 6). Fortunately, the tests commonly used for the diagnosis of influenza and SARS-CoV-2 infections have a *Sp* of almost 100% (23, 24); most of the published studies using the TNCC design have used highly specific diagnostic tests (25, 26). Considering the cardinal importance of the test *Sp* in the validity of the *VE* obtained from TNCC

design, it is prudent to use higher cut-off values in the interpretation of tests with quantitative results used to classify the cases and controls (27).

Limitations of the Study

Our study has several limitations. The simulation was limited by its simplicity; it overlooked many important factors. As an example, it was assumed that the *AR* of flu-like illness was independent of vaccination and disease status, age of people, and other variables. This was of course not the case in real life. In this way, selection of cases and controls from the subpopulation of people with flu-like illness, as it was done in the TNCC study, was practically equivalent to selecting cases and controls from the source population. That is why the traditional case-control and the TNCC designs provided almost similar results (Figure 2). If a test with a *Sp* of nearly 100% is utilized, almost no false-positive result would occur. However, if a person had had the infection of interest before vaccination, the test might remain positive for a long time, even when there is no active infection (28). Another situation that makes things complicated is usage of different vaccines (as in case of vaccination in some people against SARS-CoV-2). More complex simulations are necessary to address these limitations and provide more realistic results. Nevertheless, in this article we just intended to provide a bird's eye view of the designs commonly used in assessing the *VE*, not to provide an in-depth review of these methods.

In a nutshell, the TNCC design, which has recently been frequently used in assessing the *VE*, may reduce but not eliminate the effect of all confounding variables and selection bias due to differential recall of the exposure compared with traditional case-control design (16, 29). The chief advantage of TNCC over cohort and traditional case-control studies is the fact that it does not require much resources and it can be conducted during a relatively short period of time; it can be nested in routine surveillance without any concerns regarding the validity of the estimates derived. The design is not only commonly used in assessing the *VE*, but can also be utilized to measure risks in other settings such as antibiotic resistance (30), and venous thrombosis (31), to name only a few other applications. Using logistic regression analysis or stratification, the obtained estimated *OR* can be adjusted for important confounding variables (22). Although the method seems to be easy to do, it should be used with caution as the design suffers from all the limitations mentioned for observational studies (9, 10, 13, 32, 33). The TNCC design provides a unique opportunity for the interdisciplinary collaboration between laboratory sciences and epidemiology considering the important caveats in both areas of investigation.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

1. De Serres G, Skowronski DM, Wu XW, Ambrose CS. The test-negative design: validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. Euro Surveill. 2013;18(37):20585. doi: 10.2807/1560-7917.es2013.18.37.20585.
2. Jackson ML, Rothman KJ. Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness. Vaccine. 2015;33(11):1313-6. doi: 10.1016/j.vaccine.2015.01.069. Epub 2015 Feb 7.
3. Fukushima W, Hirota Y. Basic principles of test-negative design in evaluating influenza vaccine effectiveness. Vaccine. 2017;35(36):4796-800. doi: 10.1016/j.vaccine.2017.07.003.
4. Patel MK, Bergeri I, Bresee JS, Cowling BJ, Crowcroft NS, Fahmy K, et al. Evaluation of post-introduction COVID-19 vaccine effectiveness: Summary of interim guidance of the World Health Organization. Vaccine. 2021;39(30):4013-24. doi: 10.1016/j.vaccine.2021.05.099. Epub 2021 Jun 1.
5. World Health O. Evaluation of COVID-19 vaccine effectiveness: interim guidance, 17 March 2021. Geneva: World Health Organization; 2021. Contract No.: WHO/2019-nCoV/vaccine_effectiveness/measurement/2021.1.
6. Orenstein EW, De Serres G, Haber MJ, Shay DK, Bridges CB, Gargiullo P, et al. Methodologic issues regarding the use of three observational study designs to assess influenza vaccine effectiveness. Int J Epidemiol. 2007;36(3):623-31. doi: 10.1093/ije/dym021.
7. Hoberman A, Greenberg DP, Paradise JL, Rockette HE, Lave JR, Kearney DH, et al. Effectiveness of inactivated influenza vaccine in preventing acute otitis media in young children: a randomized controlled trial. JAMA. 2003;290(12):1608-16. doi: 10.1001/jama.290.12.1608.

8. Vesa S, Kleemola M, Blomqvist S, Takala A, Kilpi T, Hovi T. Epidemiology of documented viral respiratory infections and acute otitis media in a cohort of children followed from two to twenty-four months of age. Pediatr Infect Dis J. 2001;20(6):574-81. doi: 10.1097/00006454-200106000-00006.
9. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. Lancet. 2002;359(9300):57-61. doi: 10.1016/S0140-6736(02)07283-5.
10. Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. Lancet. 2002;359(9303):341-5. doi: 10.1016/S0140-6736(02)07500-1.
11. O'Neill RT. On sample sizes to estimate the protective efficacy of a vaccine. Stat Med. 1988;7(12):1279-88. doi: 10.1002/sim.4780071208.
12. World Health Organization. Sample size calculator for evaluation of COVID-19 vaccine effectiveness (Excel): World Health Organization; 2021 [updated March 17, 2021] [July 1, 2022]. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-vaccine_effectiveness-measurement_tool-2021.1
13. Schulz KF, Grimes DA. Case-control studies: research in reverse. Lancet. 2002;359(9304):431-4. doi: 10.1016/S0140-6736(02)07605-5.
14. Miettinen OS. Theretical Epidemiology. Principles of Occurence Research in Medicine. Hoboken, New Jersey: John Wiley and Sons; 1985.
15. Effectiveness of vaccine against medical consultation due to laboratory-confirmed influenza: results from a sentinel physician pilot project in British Columbia, 2004-2005. Can Commun Dis Rep. 2005;31(18):181-91.
16. Sullivan SG, Tchetgen Tchetgen EJ, Cowling BJ. Theoretical Basis of the Test-Negative Study Design for Assessment of Influenza Vaccine Effectiveness. Am J Epidemiol. 2016;184(5):345-53. doi: 10.1093/aje/kww064.
17. Foppa IM, Ferdinands JM, Chaves SS, Haber MJ, Reynolds SB, Flannery B, et al. The case test-negative design for studies of the effectiveness of influenza vaccine in inpatient settings. Int J Epidemiol. 2016;45(6):2052-9. doi: 10.1093/ije/dyw022.
18. Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. Vaccine. 2013;31(17):2165-8. doi: 10.1016/j.vaccine.2013.02.053. Epub 2013 Mar 13.
19. Skowronski DM, Masaro C, Kwindt TL, Mak A, Petric M, Li Y, et al. Estimating vaccine effectiveness against laboratory-confirmed influenza using a sentinel physician network: results from the 2005-2006 season of dual A and B vaccine mismatch in Canada. Vaccine. 2007;25(15):2842-51. doi: 10.1016/j.vaccine.2006.10.002.
20. Chua H, Feng S, Lewnard JA, Sullivan SG, Blyth CC, Lipsitch M, et al. The Use of Test-negative Controls to Monitor Vaccine Effectiveness: A Systematic Review of Methodology. Epidemiology. 2020;31(1):43-64. doi: 10.1097/EDE.0000000000001116.
21. Dean NE, Hogan JW, Schnitzer ME. Covid-19 Vaccine Effectiveness and the Test-Negative Design. N Engl J Med. 2021;385(15):1431-3. doi: 10.1056/NEJMe2113151. Epub 2021 Sep 8.
22. Sullivan SG, Cowling BJ. "Crude Vaccine Effectiveness" Is a Misleading Term in Test-negative Studies of Influenza Vaccine Effectiveness. Epidemiology. 2015;26(5):e60. doi: 10.1097/EDE.0000000000000343.
23. Druce J, Tran T, Kelly H, Kaye M, Chibo D, Kostecki R, et al. Laboratory diagnosis and surveillance of human respiratory viruses by PCR in Victoria, Australia, 2002-2003. J Med Virol. 2005;75(1):122-9. doi: 10.1002/jmv.20246.
24. Habibzadeh P, Mofatteh M, Silawi M, Ghavami S, Faghihi MA. Molecular diagnostic assays for COVID-19: an overview. Crit Rev Clin Lab Sci. 2021;58(6):385-98. doi: 10.1080/10408363.2021.1884640. Epub 2021 Feb 17.
25. Centers for Disease Control and Prevention (CDC). Early estimates of seasonal influenza vaccine effectiveness--United States, January 2013. MMWR Morb Mortal Wkly Rep. 2013;62(2):32-5.
26. Kelly H, Carville K, Grant K, Jacoby P, Tran T, Barr I. Estimation of influenza vaccine effectiveness from routine surveillance data. PLoS One. 2009;4(3):e5079. doi: 10.1371/journal.pone.0005079.
27. Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: the case of tests with continuous results. Biochem Med (Zagreb). 2016;26(3):297-307. doi: 10.11613/BM.2016.034.
28. Habibzadeh P, Sajadi MM, Emami A, Karimi MH, Yadollahie M, Kucheki M, et al. Rate of re-positive RT-PCR test among patients recovered from COVID-19. Biochem Med (Zagreb). 2020;30(3):030401. doi: 10.11613/BM.2020.030401. Epub 2020 Aug 5.
29. Vandenbroucke JP, Pearce N. Test-Negative Designs: Differences and Commonalities with Other Case-Control Studies with "Other Patient" Controls. Epidemiology. 2019;30(6):838-44. doi: 10.1097/EDE.0000000000001088.
30. Lipsitch M. Measuring and interpreting associations between antibiotic use and penicillin resistance in Streptococcus pneumoniae. Clin Infect Dis. 2001;32(7):1044-54. doi: 10.1086/319604.
31. Bloemenkamp KW, Rosendaal FR, Büller HR, Helmerhorst FM, Colly LP, Vandenbroucke JP. Risk of venous thrombosis with use of current low-dose oral contraceptives is not explained by diagnostic suspicion and referral bias. Arch Intern Med. 1999;159(1):65-70. doi: 10.1001/archinte.159.1.65.
32. Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. Lancet. 2002;359(9301):145-9. doi: 10.1016/S0140-6736(02)07373-7.
33. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet. 2002;359(9302):248-52. doi: 10.1016/S0140-6736(02)07451-2.

APPENDIX

The Simulation Program in R

```
library("ggplot2")
library("ggh4x")

N <- 2000000       # Population size
v_coverage <- 0.4  # Vaccine coverage
n <- 5000          # Sample size (in each group)
rep <- 1000        # Number of repetitions
AR <- 0.15         # Attack Rate in unvaccinated
AR_ill <- 0.30     # Attack Rate of flu-like illness
se <- 0.60         # Test sensitivity
sp <- 1.00         # Test specificity
VE <- 0.70         # Vaccine effectiveness

set.seed(123)
dat <- data.frame(dis = rep(0, N), Vac = rep(0, N), ill = rep(0, N))

dat$Vac[1:round(v_coverage*N)] <- 1
dat <- dat[sample(nrow(dat)),]       # Shuffle dat
dat$ill[1:round(AR_ill*N)] <- 1

len <- length(dat[dat$Vac == 0,]$dis)
dat[dat$Vac == 0,]$dis <- ifelse(runif(len, 0, 1) < AR, 1, 0)

len <- length(dat[dat$Vac == 1,]$dis)
dat[dat$Vac == 1,]$dis <- ifelse(runif(len, 0, 1) < AR*(1-VE), 1, 0)

dat$Vac = factor(dat$Vac, labels=c("Unvaccinated", "Vaccinated"))
dat$dis = factor(dat$dis, labels=c("Undiseased", "Diseased"))
dat$ill = factor(dat$ill, labels=c("Not ill", "Ill"))

dat_ve <- data.frame(AR = NA, se = NA, sp = NA, VE = NA, CI_lo = NA, CI_hi = NA, study = NA)

ve <- rep(NA, rep)
for (se in seq(1.0, 0.6, by = -0.2)){      #-- for test sensitivity from 0.6 to 1.0
  for (sp in seq(1.0, 0.8, by = -0.025)){  #-- for test specificity from 0.8 to 1.0
    dat_dis <- dat[dat$dis == "Diseased",]
    dat_dis$test <- ifelse(runif(nrow(dat_dis), 0, 1) < se, 1, 0)

    dat_undis <- dat[dat$dis == "Undiseased",]
    dat_undis$test <- ifelse(runif(nrow(dat_undis), 0, 1) < sp, 0, 1)

    dat <- rbind(dat_dis, dat_undis)
    dat$test = factor(dat$test, labels=c("Negative", "Positive"))

#-------------------------------------------------------------------------------------------

#--- Cohort
    dat_vac <- dat[dat$Vac == "Vaccinated",]
    dat_unvac <- dat[dat$Vac == "Unvaccinated",]
    if (nrow(dat_vac) < n | nrow(dat_unvac) < n){
      print("Cohort: Sample size too high!")
    }

    for (i in 1:rep){
      #-----------         Exposed      --------------------       Unexposed
      d <- rbind(dat_vac[sample(nrow(dat_vac), n), ], dat_unvac[sample(nrow(dat_unvac), n), ])
```

```
      t <- table(d$Vac, d$test)
      #-- 1-RR
      ve[i] <- 1-(t[2, 2]/(t[2, 1] + t[2, 2])) / (t[1, 2]/(t[1, 1] + t[1, 2]))
    }
    dat_ve <- rbind(dat_ve, data.frame(AR = AR, se = se, sp = sp,
            VE = mean(ve, na.rm=TRUE), CI_lo = as.numeric(quantile(ve, 0.025, na.rm=TRUE)),
            CI_hi = as.numeric(quantile(ve, 0.975, na.rm=TRUE)), study = "Coh"))

#-------------------------------------------------------------------------------------------

#--- Traditional Case-Control
    dat_dis <- dat[dat$test == "Positive",]      #-- Cases
    dat_undis <- dat[dat$test == "Negative",]    #-- Controls
    if (nrow(dat_dis) < n | nrow(dat_undis) < n){
      print("Case-Control: Sample size too high!")
    }

    for (i in 1:rep){
      #-----------         Diseased     --------------------       Undiseased
      d <- rbind(dat_dis[sample(nrow(dat_dis), n), ], dat_undis[sample(nrow(dat_undis), n), ])
      t <- table(d$Vac, d$test)
      #-- 1-OR
      ve[i] <- 1-(t[2, 2]/t[1, 2]) / (t[2, 1]/t[1, 1])
    }
    dat_ve <- rbind(dat_ve, data.frame(AR = AR, se = se, sp = sp,
            VE = mean(ve, na.rm=TRUE), CI_lo = as.numeric(quantile(ve, 0.025, na.rm=TRUE)),
            CI_hi = as.numeric(quantile(ve, 0.975, na.rm=TRUE)), study = "CC"))

#-------------------------------------------------------------------------------------------

#--- Test-Negative Case-Control
    d_ill <- dat[dat$ill == "Ill",]             # Include only ill people
    OR <- rep(NA, rep)

    for (i in 1:rep){
      d <- d_ill[sample(nrow(d_ill), 2*n),]
      t <- table(d$Vac, d$test)
      #-- 1-OR
      ve[i] <- 1-(t[2, 2]/t[1, 2]) / (t[2, 1]/t[1, 1])
    }
    dat_ve <- rbind(dat_ve, data.frame(AR = AR, se = se, sp = sp,
            VE = mean(ve, na.rm=TRUE), CI_lo = as.numeric(quantile(ve, 0.025, na.rm=TRUE)),
            CI_hi = as.numeric(quantile(ve, 0.975, na.rm=TRUE)), study = "TNCC"))
  }
}
dat_ve <- dat_ve[-1, ]
write.csv(dat_ve, file = "VE.csv", row.names = FALSE)

#----------------- Graphs Figure 2
#dat_ve <- read.csv("VE.csv")

dat_ve$study <- factor(dat_ve$study, levels = c("Coh", "CC", "TNCC"),
                               labels = c("Cohort", "Traditional\nCase-Control",
                                          "Test-Negative\nCase-Control"))

ggplot(dat_ve, aes(x = sp, color = as.factor(se), fill = as.factor(se),
         linetype = as.factor(study))) +
  geom_hline(yintercept = 0.70, color = "gray70", linetype="dotdash") +
  geom_line(aes(y = VE), size = 0.9, alpha = 0.6) +
  geom_point(aes(y = VE), size = 2.5, alpha=0.6) +
  xlab("Test Specificity") +
  theme_classic() +
  guides(x = "axis_truncated", y = "axis_truncated") +
```

```
scale_linetype_manual(values = c("dotted", "longdash", "solid")) +
scale_y_continuous(name="Vaccine Effectiveness",
       breaks=seq(0.2, 0.8, by = 0.1)) +
theme(aspect.ratio = 6/4,
legend.position = c(0.97, 0.03), legend.key.width=unit(1.2,"cm"),
legend.key.height=unit(0.9,"cm"),
legend.justification = c("right", "bottom"),
legend.text = element_text(size = rel(0.9)),
axis.text = element_text(size = rel(1.3), color = "black"),
axis.title = element_text(size = rel(1.9))) +
guides(color=guide_legend(title="Test Sensitivity", order = 1),
       fill=guide_legend(title="Test Sensitivity", order = 1),
       linetype=guide_legend(title="Study Type", order = 2))
```