



DIGITAL ACCESS TO
SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU



HARVARD LIBRARY
Office for Scholarly Communication

Observational studies and the difficult quest for causality: Lessons from vaccine effectiveness and impact studies

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters

Citation	Lipsitch, Marc, Ayan Jha, and Lone Simonsen. 2016. Observational studies and the difficult quest for causality: Lessons from vaccine effectiveness and impact studies. <i>International Journal of Epidemiology</i> . Forthcoming.
Published Version	10.1093/ije/dyw124
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:26579867
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

Observational studies and the difficult quest for causality: Lessons from vaccine effectiveness and impact studies

Marc Lipsitch¹

Ayan Jha^{1,2}

Lone Simonsen^{3,4}

- 1. Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA**
- 2. Health Education & Research Institute, Charleston Area Medical Center, 3211 MacCorkle Avenue SE, Charleston, West Virginia 25304, USA**
- 3. Department of Global Health, George Washington University, Milken Institute School of Public Health, Washington, District of Columbia 20052, USA**

ABSTRACT

Although randomized placebo-controlled trials (RCT) are critical to establish efficacy of vaccines at the time of licensure, important remaining questions about vaccine effectiveness (VE) – used here to include individual-level measures and population-wide impact of vaccine programs -- can only be answered once the vaccine is in use, from observational studies. However, such studies are inherently at risk for bias. Using a causal framework and illustrating with examples we review newer approaches to detecting and avoiding confounding and selection bias in three major classes of observational study design: cohort, case-control, and ecological studies. Studies of influenza VE, especially in seniors, are an excellent demonstration of the challenges of detecting and reducing such bias, and so we use influenza VE as a running example. We take a fresh look at the time-trend studies often dismissed as “ecological.” Such designs are the only observational study design that can measure the overall effect of a vaccination program (indirect (herd) as well as direct effects), and are in fact already an important part of the evidence base for several vaccines currently in use. Despite the great strides towards more robust observational study designs, challenges lie ahead for evaluating best practices for achieving robust unbiased results from observational studies. This is critical for evaluation of national and global vaccine program effectiveness.

KEY MESSAGES

- * Observational studies are a key part of the evidence base for the effects of vaccines (vaccine effectiveness, VE), especially at the population level and for rare but severe outcomes of infection, such as death. Yet these study designs suffer from the risk of bias due to confounding and other factors.
- * Using the illustrative example of influenza vaccines, we show that such biases can be large and describe three strategies -- the use of negative control outcomes, the use of the test-negative design, and the use of time-trend studies -- to detect and reduce various forms of bias in observational studies.
- * Each of these strategies can be very effective but each depends on certain assumptions for validity, and each may introduce new forms of bias into the analysis.
- * Further work is needed to develop best practices for observational VE studies in the post-licensure period, possibly relying on a combination of designs that may have different forms of bias.

INTRODUCTION

Although vaccine licensure requires evidence of vaccine safety and efficacy from randomized controlled trials (RCT), many questions about vaccine effectiveness (VE) can be answered only by observational approaches after the vaccine is in use. For example, the effects of a vaccine on a rare outcome, such as mortality, can typically be studied only through observational approaches. Even within a RCT, the effects of a vaccine on a subgroup defined after randomization (and therefore not protected by randomization against confounding and selection bias) may be of interest; such subgroup analyses are observational, despite being nested within a RCT. Once a vaccine is licensed and recommended for use, certain RCTs may face ethical challenges (1), though there may remain VE questions of interest. Finally, RCTs randomized at the individual level are designed to measure only the direct protection the vaccine offers to vaccinated persons, but not the important overall effect of vaccination on disease in the population, including that achieved by indirect protection of unvaccinated people (herd immunity). Post-licensure observational studies -- specifically ecological time-trend studies comparing population-level disease burdens in the pre- and post vaccination period -- are well-suited to measuring indirect and overall effects (2-4). Such studies deliberately measure a different quantity from that measured by individual-level studies of VE, but it is one that is highly relevant to policy making.

VE may be studied using any of the three major classes of observational study design: cohort, case-control, and ecological studies. Such observational VE assessments are subject to the effects of confounding and other forms of bias. In particular, those who do and do not receive a vaccine may differ in ways that affect the chance of experiencing the outcome (e.g. mortality or infection). If so, these differences confound the measured effect of the vaccine on the outcome. In this paper, we describe approaches to identify, address and reduce the impact of such confounding in observational VE studies. We provide examples of each approach as well as a formal account, using causal directed acyclic graphs, of how each approach attempts to address this source of bias and what assumptions are required for it to be a valid approach.

Specifically, we use a causal framework to describe the typical problem of confounding in VE studies and describe three approaches that have been used to address the problem: a) the use of negative control outcomes in a cohort or case-control study, b) the use of a laboratory test-negative control group in a study of all patients tested for the infection of interest, and c) the use of an ecological time-trend design to measure indirect and overall causal effects (5). Our running example is the problem of estimating influenza VE, a recently controversial area that illustrates how concerns about bias in the evidence base arose, how these approaches to reduce or detect bias are used, and how well they address the issues we are raising. Influenza VE studies raise nearly all the types and issues of potential bias, and all the strategies we

discuss have been recently applied. While we use influenza VE as an example, we make occasional reference to other vaccines for which these strategies have been deployed. Table 1 lists some of the key substantive insights for multiple vaccines that have been gained from the strategies we discuss.

THE GENERAL PROBLEM

In observational VE studies that compare outcomes in vaccinated vs. unvaccinated persons, these groups may differ in ways that affect their risk of infection or death, for reasons other than their vaccination status. For example, those who get vaccinated may have better access to health care for geographic, financial or other reasons; may be more inclined to seek health care; or may have access to different quality of care, for example more preventive health care services. Moreover, vaccinated persons may simply tend to be healthier with high “functional status”, allowing them to physically be able to get to an influenza vaccine appointment (6). Each of these characteristics may be predictive of the outcome measured in such studies, which may be laboratory-confirmed infection, illness or mortality (cause-specific or all-cause). If not adequately measured and adjusted for in analyses, these differences may confound the association between vaccine receipt and the outcome, biasing estimates of VE.

This issue is shown in Figure 1, a causal directed acyclic graph (DAG) (7) showing the causal connections among vaccination (V), laboratory-confirmed infection (I), death (D), and various confounders (H, health care access or health status). Directed arrows connecting these nodes indicate that there is assumed to be a direct effect of one variable on another. The absence of such an arrow means that there is assumed to be no direct effect of one variable on another. By convention, no arrow is shown for the causal effect that the study is attempting to estimate, here the effect of vaccination on infection or death. Confounders of these relationships are those variables that affect whether an individual becomes vaccinated (the exposure of interest) and also, separate from the vaccine effect, whether the individual becomes infected or dies (the outcomes of interest). We wish to identify the causal effect of the vaccine on infection or death, and to estimate these effects from the observed associations between V and outcomes I or D. To do so, we must account for potential confounding by H. For a detailed introduction to how to read such diagrams, see (8).

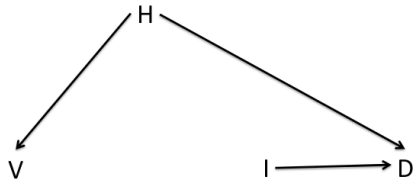


FIGURE 1: Causal structure of observational VE studies showing the possible confounding of the vaccine (V)-outcome (e.g. hospitalization or death D) relationship by health status, health-seeking behavior or other characteristics that may differ between those who do and do not receive vaccination. Any causal effect of vaccination on D must be mediated by I, infection with the pathogen against which the vaccine protects.

THREE APPROACHES TO ADDRESS BIAS IN OBSERVATIONAL VE STUDIES

1. Use of negative controls to detect confounding in studies of VE against mortality and other severe outcomes

Identifying evidence of profound bias in cohort studies of influenza VE. Influenza vaccines were developed in the 1940s for use in military populations and have since the 1960s been used mostly to vaccinate the elderly despite early concerns that immunosenescence may attenuate their effectiveness in old age (9). Once influenza vaccines were licensed and recommended for the elderly, it became ethically difficult to use a placebo-controlled randomized trial design to evaluate VE in this age group. Therefore, nearly all VE studies in the elderly have been observational, using International Classification of Disease coded diagnoses in electronic health maintenance organization (HMO) databases (for an exception see ref. (10)). These studies had consistently reported astonishing vaccine benefits -- that influenza vaccination in the elderly prevents as many as 50% of all deaths during winter (11-13). A decade ago, concerns arose that this assessment far exceeded what reasonably could be expected (6, 14, 15) and might reflect uncontrolled confounding.

Any causal effect of influenza vaccination on mortality should be evident only during influenza-epidemic periods, usually in mid-winter. Thus, any mortality advantage among vaccinees occurring in the immediate pre-epidemic months would indicate bias, likely due to confounding (6, 16). To test for the existence of bias, Jackson et al first reproduced in HMO data the original

finding of 50% winter all-cause mortality reduction, then went on to stratify the measurements according to timing (pre-, during, and after influenza). They found that the largest reduction occurred in the pre-influenza period and concluded that such studies had likely greatly exaggerated the true vaccine effect. Later studies showed that this confounding was particularly severe in studies of all-cause mortality but less of a problem when studying more specific outcome, such as hospitalization or mortality from pneumonia.

Testing for an effect of vaccination on an outcome that could not plausibly be affected by the vaccine is an example of a more general strategy that has been called the “negative control outcome” strategy. Though initially controversial, this bias-detection approach is now commonly used in studies of influenza VE (14, 17-21). Recent observational studies also use more specific endpoints (e.g., pneumonia hospitalizations instead of all cause mortality) for which VE can be more reliably estimated. The newer studies demonstrated that the VE against these endpoints was in fact low in seniors, helping to stimulate the development of more immunogenic vaccines for elderly (19).

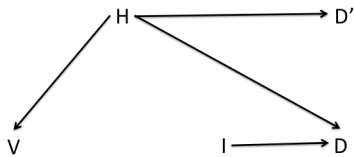
A causal account of the use of negative control outcomes to detect bias. This finding inspired formal research on the use of this strategy for bias detection in epidemiologic studies more generally. This approach has been termed a “negative control outcome” analysis, by analogy to laboratory experiments, in which researchers include “negative controls” to verify that the experimental system shows no effect when it shouldn’t (22).

Figure 2 shows a modified DAG similar to that in Figure 1, but now including a negative-control outcome, death before the influenza epidemic period (D'). By assumption, vaccination (V) has no causal effect on death outside of influenza season, but the common causes of V and D (that is, the confounders of the $V \rightarrow D$ relationship) should also be causes of D' . In other words, any association between V and D' should reflect pure confounding, in the absence of any causal pathway from V to D . Applying an analytic strategy to D' (e.g., performing a regression of the mortality outcome D' with a particular set of exposures, focusing on the estimate of the coefficient for V), and finding a significant protective effect of V , indicates that the same strategy applied to an outcome D is likely to provide an biased estimate of the causal effect of V on D . Moreover, under certain assumptions the direction of bias should be the same for the outcome of interest (D) as for the negative control (D'): in this case, the measured VE against death should be biased upwards (22).

The assumptions underlying the valid use of a negative control outcome are (i) that the exposure does not have a causal effect on the negative control outcome, either directly or indirectly, and (ii) that all confounders of the exposure-outcome relationship have the same

causal effect on the negative control outcome as they have on the outcome. Assumption 1 may be satisfied if the influenza epidemic period is accurately characterized and if influenza vaccination has no effect, directly or indirectly, on mortality that is not caused by influenza infection. Both of these are plausible premises (nonspecific effects of some vaccines on deaths not caused by the pathogen in the vaccine have been reported (23), though not to our knowledge for influenza vaccine). Assumption 2 is reasonable in the temperate regions if those factors that predict whether an individual gets vaccinated in say, early autumn, are equally predictive of mortality before influenza season as during influenza season.

FIGURE 2: As Figure 1, but showing a negative control outcome D' (for example death outside of influenza season) that has no causal connection to vaccination but shares the same confounding relationship with H as the outcome of interest D .



The use of negative controls to detect the existence and direction of bias has several advantages. Under the assumptions above, negative control outcomes can detect confounding and other forms of bias, even if the investigator has not identified the likely source of the bias ahead of time. Similarly, if one has identified a potential confounder of the relationship of interest and wants to know whether it has been adequately controlled in an analysis, the analysis can be applied to a negative control outcome to find out. Indeed, Jackson et al. did just that to test the adequacy of “standard” statistical approaches (multiple regression modeling) to controlling for health status in the elderly, and found that those had in fact been counterproductive and had exacerbated the overestimate of VE (6).

The principal limitation of the negative control outcome approach is that it does not offer a straightforward formula to correct for these biases. It can urge caution but cannot solve the problem it identifies. However, identifying the bias may be a first step toward refining the analysis to reduce it; these refined analyses may then be submitted to negative control analyses to see if the problem has been resolved. Work is underway to expand the idea of negative-controls to correct for, rather than just detect, biases (24, 25).

2. Test-negative designs

Confounding in studies of VE against clinically apparent infection. To assess VE against clinically attended infection, investigators often choose the case-control design for its ability to give results rapidly and with moderate sample sizes. As in other observational designs, vaccinated and unvaccinated persons may differ in systematic ways that may lead to confounding of the vaccine's effect. An increasingly popular, convenient and cost-saving approach to such VE studies is the use of a "test-negative" control group. While "cases" in such studies are individuals with a defined clinical syndrome who test positive for the pathogen for which the vaccine is designed, "test-negative" controls are those with the same clinical syndrome who had tested negative (1). This novel approach goes by different names and is sometimes seen as a variant of the case-control approach (26). Because it differs from a case-control study in that participants are ascertained prior to knowledge of their outcome, without a fixed ratio of outcomes in the study (27), it could be seen as a type of cohort study; indeed, arguably the first test-negative vaccine study was called an "indirect cohort" study (28). To avoid semantic confusion we simply call it the test-negative design. The motivating idea for the design is that, if a major source of confounding in the vaccination-disease relationship is a differential tendency to seek health care when ill between those who get vaccinated and those who do not, then limiting the analysis to those who have sought health care for a similar illness should reduce or eliminate this source of confounding.

It has over the last decades become established practice in the US, Canada and Europe to base influenza VE estimates on such test-negative studies conducted during and immediately after each influenza season. The "cases" and "test-negative controls" both sought care for an influenza-like illness, typically defined as fever plus respiratory symptoms (17, 26, 29-31), and the test-negative controls are those who test negative on the real-time reverse-transcriptase polymerase chain reaction (rRT-PCR) test for influenza virus.

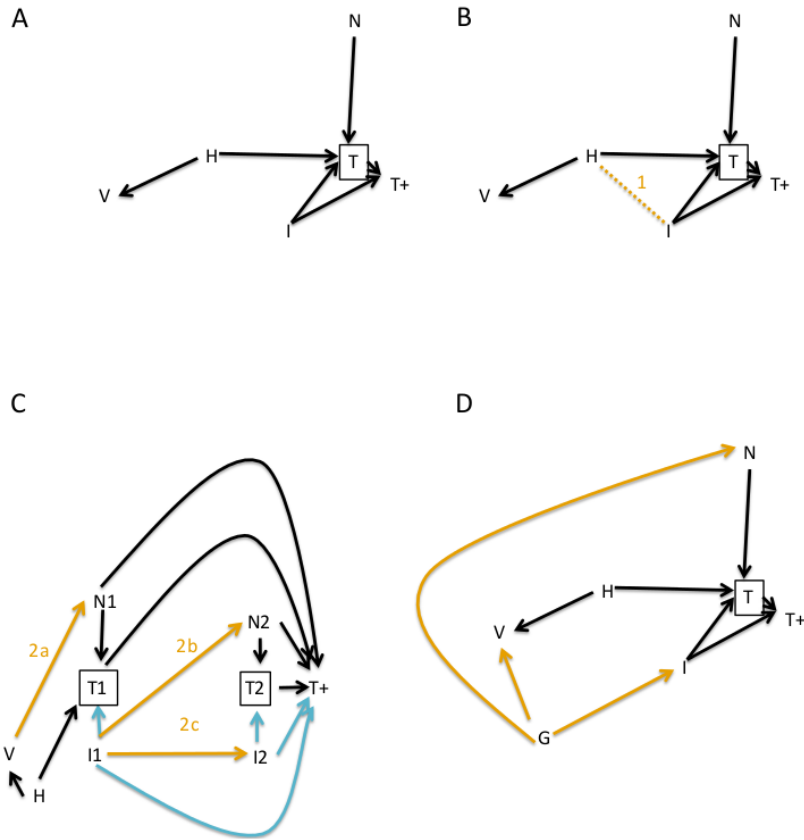


FIGURE 3: Causal diagrams of the test-negative design. Here, the outcome of interest is laboratory-confirmed clinically attended influenza infection (T^+). **A.** As before, health seeking behavior or other characteristics H may differ between those who do and do not receive vaccination, and may affect the probability of receiving a test for influenza (T) through other confounding pathways ($H \rightarrow T$). The test-negative design conditions on T by including only those who are tested, and thereby blocks this confounding effect. **B.** However, this conditioning creates selection bias, a noncausal association (orange line 1) between health-seeking behavior and infection, by conditioning on their common effect, biasing the $V \rightarrow T^+$ association. **C.** The test-negative design in greater detail, including two time periods (e.g. consecutive weeks) in which individuals are enrolled. The study is intended to measure protection by the pathways shown in green. Bias may occur if (arrow 2a) vaccination has a short-term, nonspecific effect on other infections N , or (arrow 2b) if influenza is temporarily protective against other infections N , or (arrow 2c) if influenza is protective against influenza later in the season. **D.** The test-negative design does not protect (nor does the “classic” case-control design) against confounding effects in which some factor G (e.g. being a health care worker) is a common cause of both getting vaccinated and getting infected given vaccination status.

A causal account of the test-negative design

The use of test-negative controls means that we are not studying the relationship between vaccine (V in Figure 3A) and influenza infection (I), but the relationship between vaccine and testing positive for influenza (T+) among those who are tested (T). In our expanded causal diagram for this design, we also include other factors that affect whether an individual gets tested for influenza (N), such as other infections. Figure 3A illustrates that by conditioning on testing, we block the confounding effect of health-care seeking behavior (H, whether measured or not) on the vaccination-influenza-positive relationship. In the language of causal DAGs, we block the back-door path between vaccination and the outcome $V \leftarrow H \rightarrow T+$.

The advantage of avoiding this form of confounding is balanced by several disadvantages, which can also be understood in the context of a causal diagram and are shown by the numbered arrows in Figures 3B-D. We have added further arrows (shaded orange for clarity), each representing one of the concerns described below.

Concern #1: Selection bias introduced by limiting consideration to those tested. Both influenza infection and health care seeking behavior, by assumption, influence the probability of being tested for influenza. The use of test-negative controls is equivalent to conditioning the analysis on testing, which is a common consequence of both health-seeking behavior and influenza infection. This induces a correlation among the tested persons between their health seeking behavior and influenza positivity, which induces selection bias in the estimate of VE (32). This form of selection bias or "collider bias", correlation between the two causes of a common effect when conditioning on that common effect is shown as a dashed line (Line 1, Figure 3B).

Concern #2: Bias due to measurement of effects other than those directly through preventing influenza. The scientific goal of VE studies is to measure the direct effect of influenza vaccination on preventing clinically attended, laboratory-confirmed influenza infection. If the vaccine has other effects that can mimic or mask this direct effect, then the effect will be overestimated or underestimated, respectively.

Figure 3C illustrates these biases, expanding Figure 3A to show that VE studies take place over time – in the case of influenza VE studies, usually over a full season. For simplicity we show only two time points, 1 and 2, which might be consecutive weeks during the season. The arrows shown in cyan indicate the pathways through which the test-negative design is meant to estimate VE, through the effect on preventing influenza infection. If a person is less likely to become influenza-infected, s/he is less likely to seek testing and less likely, if tested, to test

positive for influenza. The orange arrows show three pathways by which this effect could be biased.

The first such potential source of bias occurs if the seasonal influenza vaccine has an effect (for example protective) on non-influenza conditions that present with influenza-like symptoms. In this case vaccination could cause someone to be less likely to be a control than a typical member of the general population (Arrow 2a). Typically, it is assumed that this cannot happen because influenza vaccines confer antigen-specific immunity that should be protective only against influenza. Yet at least some influenza vaccines induce innate, non-specific antiviral immune responses (33) that could reduce the risk of other viral infections for a short period of time after immunization, possibly an example of a more general phenomenon (34). Graphically, this corresponds to a pathway by which vaccination increases the probability of testing positive among those tested, not via any effect on influenza, but by reducing the frequency of other conditions that to testing, $V \rightarrow N \rightarrow T+$. If this occurs, the test-negative design will be biased to show less protective effect *against influenza* because some of the benefit is hidden by protection against other infections.

The second such potential source of bias occurs if influenza infection itself induces nonspecific protection against later acquiring non-influenza infections (Arrow 2b). In this case, the vaccine's protection against influenza in any given week could increase the risk of non-influenza respiratory infection, and hence of being a test-negative participant, in a subsequent week (Arrow 2b). There is some evidence for such an effect (35). This bias will occur only if there is a true benefit of influenza vaccine against influenza infection, but will lead to an *overestimate of this protection* because vaccinated individuals will be less likely to appear in the control group than in the general population. Graphically, if the vaccine has some effectiveness against influenza infection, there will be a path $V \rightarrow I1 \rightarrow N2 \rightarrow T+$ that includes an effect the study is not designed to measure.

The third such potential source of bias occurs because people who get influenza once in a season are usually immune to getting it again that season (36) (Arrow 2c). Therefore as the influenza season progresses, if the vaccine is effective, persons at high risk of influenza infection will be removed more rapidly from the unvaccinated than the vaccinated group, and the apparent effectiveness will decline over the course of the season (37). Thus the within-season immunity will *attenuate VE estimates over the course of the season*, if the vaccine is effective. This appears to have happened in the 2012-3 influenza season in the US, where VE estimates declined over the course of the season (38-40). This source of bias, unlike the previous two, occurs even with an "ordinary" case-control design, as it does not depend on the choice of controls but rather on the fact that influenza is an immunizing infection. The graphical representation of this bias is the path $V \rightarrow I1 \rightarrow I2 \rightarrow T+$.

Concern #3: residual healthy vaccinee effects. The use of test-negative controls will eliminate confounding that arises because vaccinated, influenza infected people have a different probability of getting tested from unvaccinated, influenza-infected people -- for example because they are greater users of health care services. However, those who get vaccinated may also be more likely to get influenza than those who don't, for example because of their age or geography. We indicate such confounding with a new factor G in Figure 3D, which plausibly affects risk of both influenza and other infections. If these other factors are not separately controlled for, the use of test-negative controls will not remove their confounding effects. This problem arises whether conventional controls or test-negative controls are used, so is not unique to the test-negative design; we mention it merely to highlight that it is not solved by the test-negative design.

Concern #4: misclassification of false-negative cases as controls. Use of any diagnostic test for the identification of cases and controls has the potential of misclassification. Imperfect sensitivity of the diagnostic test leads to false negatives, so that some who should have been cases become controls. Imperfect diagnostic test specificity allows true influenza-negative individuals to become misclassified as cases. Either form of measurement error may lead to attenuation of VE estimates. A recent simulation study, however, indicates that with realistic values of VE and currently available influenza tests, the magnitude of this bias is negligible (41).

Notwithstanding the possibility of all of these forms of bias, test-negative studies have reported plausible results, including very low or no effectiveness of severely mismatched vaccine formulations for certain seasons such as 2014-15 (42-44) and for elderly in multiple seasons (30, 40). The test-negative study design was used in a study that argued that receipt of seasonal influenza vaccine was associated with increased risk of H1N1pdm morbidity during the 2009 pandemic period in Canada (17) It was also used to demonstrate that live attenuated vaccine was less effective than TIV in adults (45) and children (46). Finally, IMOVE, a large multi-country European study, routinely uses the case-negative approach to produce one pooled influenza VE estimate during each season across Europe (47).

3. Time-Trend VE studies

Time-trend study designs offer another approach to avoid some forms of confounding in VE studies. These studies, which are ecological in design, do not compare the incidence of infection or other outcomes in vaccinated vs. unvaccinated persons, so they are not subject to the kinds of confounding by health status or health-seeking behavior described in the previous sections. Instead, they estimate the reduction in an outcome following vaccine introduction in a population, where the outcome is measured in individuals (who either die or survive) but summed over the whole population. Thus, such studies estimate the effect of a change in the proportion vaccinated in a population on the frequency of that outcome in the population – a causal quantity of policy relevance. Such studies have been widely used to demonstrate and quantify the effect of vaccines including pneumococcal conjugate vaccines (2, 29, 48, 49), *Haemophilus influenzae* type b conjugate vaccines, influenza vaccines (3, 15, 50), and rotavirus vaccine (51).

Time-trend VE studies aim to measure a different quantity from the VE that is estimated in individual-level studies such as cohort and case-control studies. The impact of changing coverage of a vaccine in a population is of considerable interest to policy makers, who would like to know the number of cases that would be prevented under a vaccination program targeting a particular segment of the population. Among observational designs, an ecological study is uniquely appropriate to measure exactly this quantity. For contagious diseases, the impact of vaccinating a substantial segment of the population will be a combination of the protection to vaccinees, and indirect (“herd immunity”) protection for those who are not immunized. Here the terms direct and indirect are used in their infectious disease sense, which differs from that typical in the causal inference literature (52). In infectious diseases we use indirect protection to refer to the phenomenon that preventing infection in the vaccinated members of the population reduces infection risk to those who are unvaccinated (53). In Figure 4, we provide a modified causal diagram that includes the same elements as Figure 1, with the addition of year (Y), vaccine coverage (C). We do not consider more than one year, reflecting the assumption that vaccination and infection dynamics in one year do not affect those in future years. This is a simplification but a commonly-made one for influenza; relaxing this assumption would involve complexity beyond our scope. We show the hypothesized routes of vaccine coverage effect on infection: direct effects on the vaccinees ($C \rightarrow V \rightarrow I$) and indirect effects through the vaccine status of other members of the community on the infection risk of an individual ($C \rightarrow V' \rightarrow I$). Often the outcome measured is mortality or hospitalization (D).

By estimating the causal effect of coverage on the total number of outcomes of interest (for example, pneumonia deaths when focusing on influenza vaccine), a time-trend study

circumvents the confounding by individual health-seeking behaviors H that affects other study designs; H is by assumption not a confounder of the relationship between coverage (which can be interpreted as the average probability of vaccination in the population) and the outcome.

Time-trend studies may be used to study the overall effect of vaccination, that is the reduction in average risk for a person in a vaccinated population with a given level of coverage, compared to a person in an unvaccinated population (54, 55). Moreover, it may be limited to age groups that are not eligible for vaccination, studying the indirect effect of the vaccine, which is the impact of increased vaccine coverage in one age group on outcomes in another (3). This approach has also been used to study the indirect (herd immunity) benefits of infant PCV vaccine use as a way to protect adults (29, 49, 55, 56).

For influenza, researchers in several countries used a time-trend study to estimate the reduction in pneumonia and influenza mortality in elderly over a few decades as vaccination coverage in this age group increased from marginal to over 65% -- but failed to see any downward trends (15, 57). It was these studies that led to a re-examination of the evidence base from other observational studies that had reported astonishing mortality benefits (14). Ultimately all of these efforts led to the understanding that more immunogenic vaccine formulations may be needed for seniors, and the renewed interest in pursuing strategies for influenza control, such as the vaccination of “transmitters” including children to achieve herd protection of elderly (58).

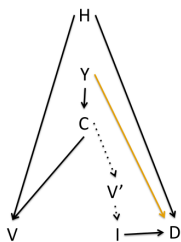


FIGURE 4: In a time-trend study of vaccine effects, year (Y) affects vaccine coverage C, which can be interpreted as each individual’s risk of being vaccinated V. Coverage also affects the vaccination status of contacts V’, which affects (through herd immunity) an individual’s infection risk. The association between coverage and the outcome D (typically a severe one, such as pneumonia or death) will be causal if year is affecting the outcome only through vaccine coverage and not through other paths such as development – that is, if the orange arrow is absent.

Conditions for the validity of time-trend VE studies

Ecological study designs have been justifiably criticized when the question of interest is how a particular exposure, such as vaccination, affects an individual’s risk of an outcome, such as death. The term “ecological fallacy” is used to describe the problem with ecological studies that

aim to estimate such individual-level quantities. However, this line of criticism does not apply to ecological VE studies, properly interpreted, because they do not aim to make inferences about individual vaccination and individual risk, but rather about the impact of changing vaccine coverage on the risk of a group of people.

Thus an ecological study comparing outcomes in populations with different levels of coverage may be considered as the non-randomized analog of a cluster-randomized trial. An ecological VE study may compare incidence within a single population in different years where vaccination coverage was different, or across different populations with different levels of coverage. In this respect, an ecologic study design is the proper observational study design to estimate overall vaccine effects. Like other observational studies, this design may suffer from confounding when populations that have different levels of vaccine coverage also differ in incidence of the outcome for reasons separate from the level of vaccine coverage. These confounding factors may include the age-distribution of the population, economy/development (which may influence both vaccine coverage and risk of severe outcomes of an illness), the recent incidence of the disease in the population (which might affect the prevalence of natural immunity to the disease), or the extent of co-circulation of other agents that can affect the outcome (for example, bacterial infections that can lead to complications of influenza infection). Such factors may confound estimates of VE, because they imply that year (in our example) has an effect on the outcome through some variable other than vaccine coverage (there may be an arrow from Y directly to D). In such a situation, it may be possible anticipate the direction of the bias induced by such confounding, or even to measure and adjust for it; for example, population aging or comparison of communities with different age structures could be improved by using age-standardized mortality rather than crude mortality (15). Likewise, adjusting for the expected reduction in severe infection outcomes as a result of socio-economic development is needed when studies are set in lower and middle income countries (59). Practically, if baseline data on the outcome are available for a period before the vaccine is introduced, evidence of a trend prior to vaccine introduction could be accounted for analytically, though there is no guarantee that extrapolating the pre-vaccine trend will accurately capture what would have happened without the vaccine.

These considerations also urge caution about the external validity of a time-trend study of VE. The proportion of a nonspecific outcome preventable with a particular vaccine will depend on the mix of causes of that outcome. For example, a vaccine against any individual pathogen that causes pneumonia will likely prevent a changing proportion of childhood pneumonia deaths as a population goes through epidemiologic transition and profoundly lower background childhood mortality (59).

An important consideration of time-trend VE studies is that the disease or mortality outcome under study must be sufficiently specific so that the vaccine's impact on it is likely to be measurable. For most vaccines, all-cause mortality is so nonspecific that the true causal effect of an effective vaccine would be expected to be only marginal, making a true effect hard to discern. Moreover, should other non-vaccine changes occur they are more likely to cause confounding, resulting in misattribution to the vaccine of other causal factors affecting general mortality.

DISCUSSION

Global vaccine efforts continue to expand thanks to the Expanded Program on Immunization (EPI) (60) and the continuous arrival of new vaccines (rotavirus, HPV, PCV, and others). Given the need for ongoing evaluation of VE of all these vaccines as they are introduced in new geographic areas and in new population segments, exclusive reliance on evidence from randomized trials is impractical for ethical, logistical and economic reasons, and is unsatisfactory because severe outcomes of most interest for policy purposes are rarely studied. Therefore the evidence base is typically built on post-introduction (phase IV) observational studies, studies characterized by a risk of bias due to confounding when differences between vaccinated and unvaccinated persons are not adequately measured and adjusted for. Many such observational studies are performed each year, but there has been no systematic assessment of the conditions under which particular study designs are to be preferred, or how disparate results should be compared. Toward that end, we have presented an overview of three of the most important approaches to improving the reliability of observational studies for estimating VE. Using the framework of causal inference and the tool of causal DAGs, we have identified conditions under which each of these study designs is valid, and thereby conditions under which they may be biased. Importantly, the existence of a possible bias does not imply that this bias will be large relative to the effect being estimated. As a corollary, the number of potential biases identified for a particular design is not an indicator of the design's desirability, as these biases may be of trivial magnitude in particular cases. Some work has been done, but more work is needed, to estimate the likely magnitude of these biases (26, 27, 41, 61-63). Such work is important because of the high reliance on such studies to provide of the evidence base used to set vaccine policy..

Failure to account for such biases has led to large errors in the estimates of vaccine effects in observational studies, and that these errors may be amplified as they gain authority and apparent precision through the process of systematic review and meta-analysis. A now well established example of this comes from influenza VE literature, in which most of the published

observational studies up to 2005 or so were afflicted with unadjusted confounding that had led to profound overestimation of vaccine benefits in seniors (14). That flawed evidence base had suggested that 1 death could be saved by vaccinating 150-300 elderly - when in fact seniors respond poorly to influenza vaccine. Once the cohort study designs had been improved and bias detection strategies incorporated (17, 18, 62) earlier reports of astonishing mortality savings were replaced with the insight that these studies suffered from confounding that led to dramatic VE overestimation. This finding was reinforced by evidence from a trend study of modest effects at the population level (15). We have discussed three approaches to avoiding or circumventing such bias in observational studies -- but each involves tradeoffs. Characteristics of the three approaches are summarized in Table 2.

Firstly, the use of negative controls in an observational study can be a powerful way to detect bias, but only if the relationship between confounders and the negative control outcome closely resembles the relationship with the outcome of interest, and if there are no other sources of association between vaccination and the negative control outcome. The use of pre-influenza mortality as a negative control outcome is particularly suitable because there is strong biological plausibility to these assumptions, but in other situations (in particular for non-seasonal diseases) the available negative control outcomes may be more open to question. Further efforts are needed to define operationally how negative control outcomes can best be selected. Meanwhile, we can perhaps think of the negative control outcomes as “sensitivity analyses”; more than typical sensitivity analyses these are meant to add biological meaning to the exploration of robustness of the VE findings.

Secondly, the test-negative approach is now widely used to provide timely and accurate VE estimates from laboratory-based surveillance data. The upside of test-negative studies is that they can eliminate important confounding effects of health-seeking behavior, but the downside is that they do so at the cost of risk of selection bias that leads to a non-causal association between health-care seeking behavior and infection with the pathogen of interest. Moreover, this design requires that neither the infection of interest, nor the vaccine of interest, affects the probability of other infections that may lead someone to be a test-negative control. These assumptions are questionable in many circumstances, though the evidence is limited for most infections. We suspect that the downsides of the test-negative approach will often be outweighed by the advantages of avoiding confounding by health-care seeking behavior, but additional simulation studies may help to throw light on the magnitude of potential biases. Meanwhile, it will often be possible to include both test-negative controls and other control groups that have different properties. Consistent results from such comparisons will be reassuring.

Thirdly, time-trend studies are often used to demonstrate benefits of vaccine programs. The advantages are that these are by design free of confounding due to individual differences by studying rates in the whole population, and uniquely address the important question of the overall population-level effect of vaccine coverage (different from case-control and cohort studies that can only measure the direct VE). The main disadvantage is that these studies are prone to their own forms of confounding, if other factors besides vaccine coverage that affect the outcome are changing over time along with vaccine coverage. Such confounding, if measured, may be demonstrated by the use of negative control outcomes, and adjusted for in various ways. Despite being somewhat unfairly stigmatized for being “ecological”, this class of studies is appropriately prominent in the evidence base because of the importance of measuring population-level effects, especially on outcomes that are too rare to measure in most RCTs.

The approaches to address bias in observational studies described here can be combined or blended. For example, time-trend studies were used to investigate the impact of pneumococcal conjugate vaccine on respiratory hospitalizations, and investigators used various negative control outcomes, such as dehydration (64) and urinary tract infection (49) hospitalizations, to rule out other population-level trends that could have affected hospitalization rates in general.

In another example, several VE case-control studies for cholera vaccines have used confirmed cholera diarrhea as cases and chose as controls persons who did not seek treatment for diarrhea – a conventional design (65-67). They then did a parallel “bias-indicator” analysis using those who sought treatment for diarrhea but did not test positive for cholera as cases and conventional controls, showing that the study design did not find measurable VE for a non-cholera outcome. This negative control outcome approach has an obvious similarity to the test-negative design.

Considering importance given of positive evidence from trends studies (4, 49, 64, 68, 69) in shaping vaccine policy, and given that the time-trend study design is the only one to capture the overall benefits of a vaccine program, it is critical that much effort goes into the identification of strategies for bias detection and bias-elimination and to the validation of findings and elimination of any other explanation than vaccine for observed reductions in an outcome. Unfortunately, the literature at this point is not coordinated with respect to bias-detection and trends adjustment strategies. Furthermore, more effort is needed to assess the extent of bias favoring publication of papers with expected or positive results.

We have not comprehensively drawn attention to every source of bias in each of the three approaches considered here, but have concentrated on those that are particular either to the

design itself or to the class of VE studies. These different kinds of study design approaches have different strengths and limitations, such that consistency of evidence from more than one of these may be more compelling than repeated, consistent findings from one study design (17). This may suggest that instead of the classical “hierarchy of evidence quality” the strength of evidence on a question should be evaluated by the consistency of findings across study designs (70, 71).

In conclusion, challenges lie ahead for evaluating best practices for achieving robust unbiased results from observational VE studies. Short of enormous cluster-randomized trials at the time of introduction of a new vaccine, these will form the major part of the evidence base used to evaluate national and global vaccine program effectiveness. Great strides have been made in recent years, but this is an unfinished agenda.

ACKNOWLEDGMENTS

We thank M Hernan for helpful discussions.

Conflict of interest: ML has received research funding through his institution for vaccine modeling studies from Pfizer and PATH Vaccine Solutions, and honoraria from Pfizer (donated to charity) for consulting on vaccine issues. LS is a co-owner of Sage Analytica, a small research consultancy which has done contracting work for Pfizer, GSK and AstraZeneca on vaccine evaluation modeling. AJ declares no COI.

REFERENCES

1. Rid A, Saxena A, Baqui AH, Bhan A, Bines J, Bouesseau MC, et al. Placebo use in vaccine trials: recommendations of a WHO expert panel. *Vaccine*. 2014;32(37):4708-12. Epub 2014/04/29. doi: 10.1016/j.vaccine.2014.04.022. PubMed PMID: 24768580; PubMed Central PMCID: PMC4157320.
2. Pilishvili T, Lexau C, Farley MM, Hadler J, Harrison LH, Bennett NM, et al. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. *J Infect Dis*. 2010;201(1):32-41. Epub 2009/12/02. doi: 10.1086/648593. PubMed PMID: 19947881.
3. Peltola H, Aavitsland P, Hansen KG, Jonsdottir KE, Nokleby H, Romanus V. Perspective: a five-country analysis of the impact of four different *Haemophilus influenzae* type b conjugates and vaccination strategies in Scandinavia. *J Infect Dis*. 1999;179(1):223-9. Epub 1998/12/08. doi: 10.1086/314535. PubMed PMID: 9841843.
4. Richardson V, Hernandez-Pichardo J, Quintanar-Solares M, Esparza-Aguilar M, Johnson B, Gomez-Altamirano CM, et al. Effect of rotavirus vaccination on death from childhood diarrhea in Mexico. *N Engl J Med*. 2010;362(4):299-305. Epub 2010/01/29. doi: 10.1056/NEJMoa0905211. PubMed PMID: 20107215.
5. Halloran ME, Longini IM, Jr., Struchiner CJ. Design and interpretation of vaccine field studies. *Epidemiol Rev*. 1999;21(1):73-88. Epub 1999/10/16. PubMed PMID: 10520474.
6. Jackson LA, Nelson JC, Benson P, Neuzil KM, Reid RJ, Psaty BM, et al. Functional status is a confounder of the association of influenza vaccine and risk of all cause mortality in seniors. *Int J Epidemiol*. 2006;35(2):345-52. Epub 2005/12/22. doi: 10.1093/ije/dyi275. PubMed PMID: 16368724.
7. Hernán MA, Robins JM. Causal Inference. Harvard T. H. Chan School of Public Health 2015. Available from: <http://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>.
8. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10(1):37-48. Epub 1999/01/15. PubMed PMID: 9888278.
9. Langmuir AD, Henderson DA, Serfling RE. THE EPIDEMIOLOGICAL BASIS FOR THE CONTROL OF INFLUENZA. *Am J Public Health Nations Health*. 1964;54:563-71. Epub 1964/04/01. PubMed PMID: 14136320; PubMed Central PMCID: PMC1254817.
10. Govaert TM, Thijs CT, Masurel N, Sprenger MJ, Dinant GJ, Knottnerus JA. The efficacy of influenza vaccination in elderly individuals. A randomized double-blind placebo-controlled trial. *JAMA*. 1994;272(21):1661-5. Epub 1994/12/07. PubMed PMID: 7966893.
11. Nichol KL, Nordin JD, Nelson DB, Mullooly JP, Hak E. Effectiveness of influenza vaccine in the community-dwelling elderly. *N Engl J Med*. 2007;357(14):1373-81. Epub 2007/10/05. doi: 10.1056/NEJMoa070844. PubMed PMID: 17914038.
12. Poland GA. If you could halve the mortality rate, would you do it? *Clin Infect Dis*. 2002;35(4):378-80. Epub 2002/07/30. doi: 10.1086/341404. PubMed PMID: 12145719.
13. Nordin J, Mullooly J, Poblete S, Strikas R, Petrucci R, Wei F, et al. Influenza vaccine effectiveness in preventing hospitalizations and deaths in persons 65 years or older in Minnesota, New York, and Oregon: data from 3 health plans. *J Infect Dis*. 2001;184(6):665-70. Epub 2001/08/23. doi: 10.1086/323085. PubMed PMID: 11517426.
14. Simonsen L, Taylor RJ, Viboud C, Miller MA, Jackson LA. Mortality benefits of influenza vaccination in elderly people: an ongoing controversy. *Lancet Infect Dis*. 2007;7(10):658-66. Epub 2007/09/28. doi: 10.1016/s1473-3099(07)70236-0. PubMed PMID: 17897608.
15. Simonsen L, Reichert TA, Viboud C, Blackwelder WC, Taylor RJ, Miller MA. Impact of influenza vaccination on seasonal mortality in the US elderly population. *Arch Intern Med*.

- 2005;165(3):265-72. Epub 2005/02/16. doi: 10.1001/archinte.165.3.265. PubMed PMID: 15710788.
16. Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. *Int J Epidemiol* 2006;35(2):337 - 44.
 17. Skowronski DM, De Serres G, Crowcroft NS, Janjua NZ, Boulianne N, Hottes TS, et al. Association between the 2008-09 seasonal influenza vaccine and pandemic H1N1 illness during Spring-Summer 2009: four observational studies from Canada. *PLoS Med*. 2010;7(4):e1000258. Epub 2010/04/14. doi: 10.1371/journal.pmed.1000258. PubMed PMID: 20386731; PubMed Central PMCID: PMC2850386.
 18. Fireman B, Lee J, Lewis N, Bembom O, van der Laan M, Baxter R. Influenza vaccination and mortality: differentiating vaccine effects from bias. *Am J Epidemiol*. 2009;170(5):650-6. Epub 2009/07/25. doi: 10.1093/aje/kwp173. PubMed PMID: 19625341; PubMed Central PMCID: PMC2728831.
 19. Jackson ML, Nelson JC, Weiss NS, Neuzil KM, Barlow W, Jackson LA. Influenza vaccination and risk of community-acquired pneumonia in immunocompetent elderly people: a population-based, nested case-control study. *Lancet*. 2008;372(9636):398-405. Epub 2008/08/05. doi: 10.1016/s0140-6736(08)61160-5. PubMed PMID: 18675690.
 20. Jackson ML, Yu O, Nelson JC, Naleway A, Belongia EA, Baxter R, et al. Further evidence for bias in observational studies of influenza vaccine effectiveness: the 2009 influenza A(H1N1) pandemic. *Am J Epidemiol*. 2013;178(8):1327-36. Epub 2013/08/28. doi: 10.1093/aje/kwt124. PubMed PMID: 23978527.
 21. Baxter R, Lee J, Fireman B. Evidence of bias in studies of influenza vaccine effectiveness in elderly patients. *J Infect Dis*. 2010;201(2):186-9. Epub 2009/12/10. doi: 10.1086/649568. PubMed PMID: 19995265.
 22. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21(3):383-8. Epub 2010/03/26. doi: 10.1097/EDE.0b013e3181d61eeb. PubMed PMID: 20335814; PubMed Central PMCID: PMC3053408.
 23. Sankoh O, Welaga P, Debpuur C, Zandoh C, Gyaase S, Poma MA, et al. The non-specific effects of vaccines and other childhood interventions: the contribution of INDEPTH Health and Demographic Surveillance Systems. *Int J Epidemiol*. 2014;43(3):645-53. Epub 2014/06/13. doi: 10.1093/ije/dyu101. PubMed PMID: 24920644; PubMed Central PMCID: PMC4052142.
 24. Richardson DB, Keil AP, Tchetgen Tchetgen E, Cooper G. Negative Control Outcomes and the Analysis of Standardized Mortality Ratios. *Epidemiology*. 2015;26(5):727-32. Epub 2015/07/15. doi: 10.1097/ede.0000000000000353. PubMed PMID: 26172862.
 25. Tchetgen Tchetgen E. The control outcome calibration approach for causal inference with unobserved confounding. *Am J Epidemiol*. 2014;179(5):633-40. Epub 2013/12/24. doi: 10.1093/aje/kwt303. PubMed PMID: 24363326; PubMed Central PMCID: PMC3927977.
 26. 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). Epub 2013/10/02. PubMed PMID: 24079398.
 27. Foppa IM, Haber M, Ferdinands JM, Shay DK. The case test-negative design for studies of the effectiveness of influenza vaccine. *Vaccine*. 2013;31(30):3104-9. Epub 2013/04/30. doi: 10.1016/j.vaccine.2013.04.026. PubMed PMID: 23624093.
 28. Broome CV, Facklam RR, Fraser DW. Pneumococcal disease after pneumococcal vaccination: an alternative method to estimate the efficacy of pneumococcal vaccine. *N Engl J Med*. 1980;303(10):549-52. Epub 1980/09/04. doi: 10.1056/nejm198009043031003. PubMed PMID: 6995835.

29. De Serres G, Pilishvili T, Link-Gelles R, Reingold A, Gershman K, Petit S, et al. Use of surveillance data to estimate the effectiveness of the 7-valent conjugate pneumococcal vaccine in children less than 5 years of age over a 9 year period. *Vaccine*. 2012;30(27):4067-72. Epub 2012/04/25. doi: 10.1016/j.vaccine.2012.04.017. PubMed PMID: 22525797.
30. Belongia EA, Kieke BA, Donahue JG, Coleman LA, Irving SA, Meece JK, et al. Influenza vaccine effectiveness in Wisconsin during the 2007-08 season: comparison of interim and final results. *Vaccine*. 2011;29(38):6558-63. Epub 2011/07/20. doi: 10.1016/j.vaccine.2011.07.002. PubMed PMID: 21767593.
31. Kissling E, Valenciano M, Cohen JM, Oroszi B, Barret AS, Rizzo C, et al. I-MOVE multi-centre case control study 2010-11: overall and stratified estimates of influenza vaccine effectiveness in Europe. *PLoS One*. 2011;6(11):e27622. Epub 2011/11/24. doi: 10.1371/journal.pone.0027622. PubMed PMID: 22110695; PubMed Central PMCID: PMC3216983.
32. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615-25. Epub 2004/08/17. PubMed PMID: 15308962.
33. Chua BY, Wong CY, Mifsud EJ, Edenborough KM, Sekiya T, Tan AC, et al. Inactivated Influenza Vaccine That Provides Rapid, Innate-Immune-System-Mediated Protection and Subsequent Long-Term Adaptive Immunity. *MBio*. 2015;6(6). Epub 2015/10/29. doi: 10.1128/mBio.01024-15. PubMed PMID: 26507227; PubMed Central PMCID: PMC4626850.
34. Benn CS, Netea MG, Selin LK, Aaby P. A small jab - a big effect: nonspecific immunomodulation by vaccines. *Trends Immunol*. 2013;34(9):431-9. Epub 2013/05/18. doi: 10.1016/j.it.2013.04.004. PubMed PMID: 23680130.
35. Cowling BJ, Nishiura H. Virus interference and estimates of influenza vaccine effectiveness from test-negative studies. *Epidemiology*. 2012;23(6):930-1. Epub 2012/10/06. doi: 10.1097/EDE.0b013e31826b300e. PubMed PMID: 23038121.
36. Goldstein E, Cobey S, Takahashi S, Miller JC, Lipsitch M. Predicting the epidemic sizes of influenza A/H1N1, A/H3N2, and B: a statistical method. *PLoS Med*. 2011;8(7):e1001051. Epub 2011/07/14. doi: 10.1371/journal.pmed.1001051. PubMed PMID: 21750666; PubMed Central PMCID: PMC3130020.
37. O'Hagan JJ, Hernan MA, Walensky RP, Lipsitch M. Apparent declining efficacy in randomized trials: examples of the Thai RV144 HIV vaccine and South African CAPRISA 004 microbicide trials. *AIDS*. 2012;26(2):123-6. Epub 2011/11/03. doi: 10.1097/QAD.0b013e32834e1ce7. PubMed PMID: 22045345; PubMed Central PMCID: PMC3319457.
38. Centers for Disease Control and Prevention. Early estimates of seasonal influenza vaccine effectiveness--United States, January 2013. *MMWR Morb Mortal Wkly Rep*. 2013;62(2):32-5. Epub 2013/01/18. PubMed PMID: 23325354.
39. Centers for Disease Control and Prevention. Interim adjusted estimates of seasonal influenza vaccine effectiveness - United States, February 2013. *MMWR Morb Mortal Wkly Rep*. 2013;62(7):119-23. Epub 2013/02/22. PubMed PMID: 23425960.
40. McLean HQ, Thompson MG, Sundaram ME, Kieke BA, Gaglani M, Murthy K, et al. Influenza vaccine effectiveness in the United States during 2012-2013: variable protection by age and virus type. *J Infect Dis*. 2015;211(10):1529-40. Epub 2014/11/20. doi: 10.1093/infdis/jiu647. PubMed PMID: 25406334; PubMed Central PMCID: PMC4407759.
41. Jackson ML, Rothman KJ. Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness. *Vaccine*. 2015;33(11):1313-6. Epub 2015/02/11. doi: 10.1016/j.vaccine.2015.01.069. PubMed PMID: 25659280.
42. Valenciano M, Kissling E, Reuss A, Jimenez-Jorge S, Horvath JK, Donnell JM, et al. The European I-MOVE Multicentre 2013-2014 Case-Control Study. Homogeneous moderate influenza vaccine effectiveness against A(H1N1)pdm09 and heterogenous results by country

- against A(H3N2). *Vaccine*. 2015;33(24):2813-22. Epub 2015/05/06. doi: 10.1016/j.vaccine.2015.04.012. PubMed PMID: 25936723.
43. Flannery B, Clippard J, Zimmerman RK, Nowalk MP, Jackson ML, Jackson LA, et al. Early estimates of seasonal influenza vaccine effectiveness - United States, January 2015. *MMWR Morb Mortal Wkly Rep*. 2015;64(1):10-5. Epub 2015/01/16. PubMed PMID: 25590680.
44. Gilca R, Skowronski DM, Douville-Fradet M, Amini R, Boulianne N, Rouleau I, et al. Mid-Season Estimates of Influenza Vaccine Effectiveness against Influenza A(H3N2) Hospitalization in the Elderly in Quebec, Canada, January 2015. *PLoS One*. 2015;10(7):e0132195. Epub 2015/07/23. doi: 10.1371/journal.pone.0132195. PubMed PMID: 26200655; PubMed Central PMCID: PMC4511737.
45. Bateman AC, Kieke BA, Irving SA, Meece JK, Shay DK, Belongia EA. Effectiveness of monovalent 2009 pandemic influenza A virus subtype H1N1 and 2010-2011 trivalent inactivated influenza vaccines in Wisconsin during the 2010-2011 influenza season. *J Infect Dis*. 2013;207(8):1262-9. Epub 2013/01/24. doi: 10.1093/infdis/jit020. PubMed PMID: 23341536.
46. Chung JR, Flannery B, Thompson MG, Gaglani M, Jackson ML, Monto AS, et al. Seasonal Effectiveness of Live Attenuated and Inactivated Influenza Vaccine. *Pediatrics*. 2016. Epub 2016/01/08. doi: 10.1542/peds.2015-3279. PubMed PMID: 26738884.
47. Kissling E, Valenciano M, Falcao J, Larrauri A, Widgren K, Pitigoi D, et al. "I-MOVE" towards monitoring seasonal and pandemic influenza vaccine effectiveness: lessons learnt from a pilot multi-centric case-control study in Europe, 2008-9. *Euro Surveill*. 2009;14(44). Epub 2009/11/28. PubMed PMID: 19941774.
48. Kyaw MH, Lynfield R, Schaffner W, Craig AS, Hadler J, Reingold A, et al. Effect of introduction of the pneumococcal conjugate vaccine on drug-resistant *Streptococcus pneumoniae*. *N Engl J Med*. 2006;354(14):1455-63. Epub 2006/04/07. doi: 10.1056/NEJMoa051642. PubMed PMID: 16598044.
49. Simonsen L, Taylor RJ, Schuck-Paim C, Lustig R, Haber M, Klugman KP. Effect of 13-valent pneumococcal conjugate vaccine on admissions to hospital 2 years after its introduction in the USA: a time series analysis. *Lancet Respir Med*. 2014;2(5):387-94. Epub 2014/05/13. doi: 10.1016/s2213-2600(14)70032-3. PubMed PMID: 24815804.
50. Reichert TA, Sugaya N, Fedson DS, Glezen WP, Simonsen L, Tashiro M. The Japanese experience with vaccinating schoolchildren against influenza. *N Engl J Med*. 2001;344(12):889-96. Epub 2001/03/22. doi: 10.1056/nejm200103223441204. PubMed PMID: 11259722.
51. Richardson V, Parashar U, Patel M. Childhood diarrhea deaths after rotavirus vaccination in Mexico. *N Engl J Med*. 2011;365(8):772-3. Epub 2011/08/26. doi: 10.1056/NEJMc1100062. PubMed PMID: 21864191.
52. VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology*. 2010;21(4):540-51. Epub 2010/05/19. doi: 10.1097/EDE.0b013e3181df191c. PubMed PMID: 20479643; PubMed Central PMCID: PMC4231822.
53. Halloran ME, Struchiner CJ, Longini IM, Jr. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. *Am J Epidemiol*. 1997;146(10):789-803. Epub 1997/12/31 23:39. PubMed PMID: 9384199.
54. Halloran ME, Haber M, Longini IM, Jr., Struchiner CJ. Direct and indirect effects in vaccine efficacy and effectiveness. *Am J Epidemiol*. 1991;133(4):323-31. Epub 1991/02/25. PubMed PMID: 1899778.
55. Simonsen L, Taylor RJ, Young-Xu Y, Haber M, May L, Klugman KP. Impact of pneumococcal conjugate vaccination of infants on pneumonia and influenza hospitalization and mortality in all age groups in the United States. *MBio*. 2011;2(1):e00309-10. Epub 2011/01/26. doi: 10.1128/mBio.00309-10. PubMed PMID: 21264063; PubMed Central PMCID: PMC3025524.

56. Griffin MR, Zhu Y, Moore MR, Whitney CG, Grijalva CG. U.S. hospitalizations for pneumonia after a decade of pneumococcal vaccination. *N Engl J Med*. 2013;369(2):155-63. Epub 2013/07/12. doi: 10.1056/NEJMoa1209165. PubMed PMID: 23841730.
57. Rizzo C, Viboud C, Montomoli E, Simonsen L, Miller MA. Influenza-related mortality in the Italian elderly: no decline associated with increasing vaccination coverage. *Vaccine*. 2006;24(42-43):6468-75. Epub 2006/08/01. doi: 10.1016/j.vaccine.2006.06.052. PubMed PMID: 16876293.
58. Halloran ME, Longini IM, Jr. Public health. Community studies for vaccinating schoolchildren against influenza. *Science*. 2006;311(5761):615-6. Epub 2006/02/04. doi: 10.1126/science.1122143. PubMed PMID: 16456066.
59. Campbell H, Nair H. Child pneumonia at a time of epidemiological transition. *Lancet Glob Health*. 2015;3(2):e65-6. Epub 2015/01/27. doi: 10.1016/s2214-109x(14)70308-0. PubMed PMID: 25617193.
60. World Health Organization. The Expanded Programme on Immunization: WHO; 2015 [cited 2015 November 4]. Available from: http://www.who.int/immunization/programmes_systems/supply_chain/benefits_of_immunization/en/.
61. Haber M, An Q, Foppa IM, Shay DK, Ferdinands JM, Orenstein WA. A probability model for evaluating the bias and precision of influenza vaccine effectiveness estimates from case-control studies. *Epidemiol Infect*. 2015;143(7):1417-26. Epub 2014/08/26. doi: 10.1017/s0950268814002179. PubMed PMID: 25147970; PubMed Central PMCID: PMC4336850.
62. Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine*. 2013;31(17):2165-8. Epub 2013/03/19. doi: 10.1016/j.vaccine.2013.02.053. PubMed PMID: 23499601.
63. 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. Epub 2007/04/04. doi: 10.1093/ije/dym021. PubMed PMID: 17403908.
64. Grijalva CG, Nuorti JP, Arbogast PG, Martin SW, Edwards KM, Griffin MR. Decline in pneumonia admissions after routine childhood immunisation with pneumococcal conjugate vaccine in the USA: a time-series analysis. *Lancet*. 2007;369(9568):1179-86. Epub 2007/04/10. doi: 10.1016/s0140-6736(07)60564-9. PubMed PMID: 17416262.
65. Ivers LC, Hilaire IJ, Teng JE, Almazor CP, Jerome JG, Ternier R, et al. Effectiveness of reactive oral cholera vaccination in rural Haiti: a case-control study and bias-indicator analysis. *Lancet Glob Health*. 2015;3(3):e162-8. Epub 2015/02/24. doi: 10.1016/s2214-109x(14)70368-7. PubMed PMID: 25701994; PubMed Central PMCID: PMC4384694.
66. Luquero FJ, Grout L, Ciglenecki I, Sakoba K, Traore B, Heile M, et al. Use of *Vibrio cholerae* vaccine in an outbreak in Guinea. *N Engl J Med*. 2014;370(22):2111-20. Epub 2014/05/30. doi: 10.1056/NEJMoa1312680. PubMed PMID: 24869721.
67. Lucas ME, Deen JL, von Seidlein L, Wang XY, Ampuero J, Puri M, et al. Effectiveness of mass oral cholera vaccination in Beira, Mozambique. *N Engl J Med*. 2005;352(8):757-67. Epub 2005/02/25. doi: 10.1056/NEJMoa043323. PubMed PMID: 15728808.
68. Whitney CG, Farley MM, Hadler J, Harrison LH, Bennett NM, Lynfield R, et al. Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *N Engl J Med*. 2003;348(18):1737-46. Epub 2003/05/02. doi: 10.1056/NEJMoa022823. PubMed PMID: 12724479.
69. Feikin DR, Kagucia EW, Loo JD, Link-Gelles R, Puhan MA, Cherian T, et al. Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. *PLoS Med*. 2013;10(9):e1001517.

- Epub 2013/10/03. doi: 10.1371/journal.pmed.1001517. PubMed PMID: 24086113; PubMed Central PMCID: PMC3782411.
70. Rosenbaum PR. How to See More in Observational Studies: Some New Quasi-Experimental Devices. *Annual Review of Statistics and Its Application*. 2015;2:21-48. doi: 10.1146/annurev-statistics-010814-020201.
 71. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA*. 1995;274(22):1800-4. Epub 1995/12/13. PubMed PMID: 7500513.
 72. Treanor JJ, Talbot HK, Ohmit SE, Coleman LA, Thompson MG, Cheng PY, et al. Effectiveness of seasonal influenza vaccines in the United States during a season with circulation of all three vaccine strains. *Clin Infect Dis*. 2012;55(7):951-9. Epub 2012/07/31. doi: 10.1093/cid/cis574. PubMed PMID: 22843783; PubMed Central PMCID: PMC3657521.
 73. Griffin MR, Monto AS, Belongia EA, Treanor JJ, Chen Q, Chen J, et al. Effectiveness of non-adjuvanted pandemic influenza A vaccines for preventing pandemic influenza acute respiratory illness visits in 4 U.S. communities. *PLoS One*. 2011;6(8):e23085. Epub 2011/08/23. doi: 10.1371/journal.pone.0023085. PubMed PMID: 21857999; PubMed Central PMCID: PMC3155536.
 74. Blyth CC, Jacoby P, Effler PV, Kelly H, Smith DW, Robins C, et al. Effectiveness of trivalent flu vaccine in healthy young children. *Pediatrics*. 2014;133(5):e1218-25. Epub 2014/04/23. doi: 10.1542/peds.2013-3707. PubMed PMID: 24753525.
 75. Hellenbrand W, Jorgensen P, Schweiger B, Falkenhorst G, Nachtnebel M, Greutelaers B, et al. Prospective hospital-based case-control study to assess the effectiveness of pandemic influenza A(H1N1)pdm09 vaccination and risk factors for hospitalization in 2009-2010 using matched hospital and test-negative controls. *BMC Infect Dis*. 2012;12:127. Epub 2012/06/02. doi: 10.1186/1471-2334-12-127. PubMed PMID: 22650369; PubMed Central PMCID: PMC3464893.
 76. Cortese MM, Immergluck LC, Held M, Jain S, Chan T, Grizas AP, et al. Effectiveness of monovalent and pentavalent rotavirus vaccine. *Pediatrics*. 2013;132(1):e25-33. Epub 2013/06/19. doi: 10.1542/peds.2012-3804. PubMed PMID: 23776114; PubMed Central PMCID: PMC4074617.
 77. Correia JB, Patel MM, Nakagomi O, Montenegro FM, Germano EM, Correia NB, et al. Effectiveness of monovalent rotavirus vaccine (Rotarix) against severe diarrhea caused by serotypically unrelated G2P[4] strains in Brazil. *J Infect Dis*. 2010;201(3):363-9. Epub 2010/01/06. doi: 10.1086/649843. PubMed PMID: 20047501.
 78. Patel M, Pedreira C, De Oliveira LH, Umana J, Tate J, Lopman B, et al. Duration of protection of pentavalent rotavirus vaccination in Nicaragua. *Pediatrics*. 2012;130(2):e365-72. Epub 2012/07/04. doi: 10.1542/peds.2011-3478. PubMed PMID: 22753550.
 79. Patel MM, Patzi M, Pastor D, Nina A, Roca Y, Alvarez L, et al. Effectiveness of monovalent rotavirus vaccine in Bolivia: case-control study. *BMJ*. 2013;346:f3726. Epub 2013/06/21. doi: 10.1136/bmj.f3726. PubMed PMID: 23783434; PubMed Central PMCID: PMC3687514.
 80. Klein NP, Bartlett J, Rowhani-Rahbar A, Fireman B, Baxter R. Waning protection after fifth dose of acellular pertussis vaccine in children. *N Engl J Med*. 2012;367(11):1012-9. Epub 2012/09/14. doi: 10.1056/NEJMoa1200850. PubMed PMID: 22970945.
 81. Baxter R, Bartlett J, Rowhani-Rahbar A, Fireman B, Klein NP. Effectiveness of pertussis vaccines for adolescents and adults: case-control study. *BMJ*. 2013;347:f4249. Epub 2013/07/23. doi: 10.1136/bmj.f4249. PubMed PMID: 23873919.
 82. Benin AL, O'Brien KL, Watt JP, Reid R, Zell ER, Katz S, et al. Effectiveness of the 23-valent polysaccharide vaccine against invasive pneumococcal disease in Navajo adults. *J Infect Dis*. 2003;188(1):81-9. Epub 2003/06/26. doi: 10.1086/375782. PubMed PMID: 12825175.

83. Doshi RH, Mukadi P, Shidi C, Mulumba A, Hoff NA, Gerber S, et al. Field evaluation of measles vaccine effectiveness among children in the Democratic Republic of Congo. *Vaccine*. 2015;33(29):3407-14. Epub 2015/05/06. doi: 10.1016/j.vaccine.2015.04.067. PubMed PMID: 25937449.
84. Pilishvili T, Bennett NM. Pneumococcal disease prevention among adults: Strategies for the use of pneumococcal vaccines. *Vaccine*. 2015. Epub 2015/06/28. doi: 10.1016/j.vaccine.2015.05.102. PubMed PMID: 26116257.

TABLE 1: Substantive insights from published literature using the three approaches described in this paper

Substantive insights from negative-control approaches

- Showed that the vaccine-preventable burden of influenza mortality has been overestimated due to confounding (6, 14, 15)
- Added confidence to several studies of cholera VE in outbreaks by showing that VE was estimated to be not statistically different from zero for non-cholera diarrhea (65-67)

Substantive insights from test-negative designs

- Measured and compared VE of pandemic and seasonal vaccines in various seasons documenting variation across seasons and sometimes across vaccines within a season in studies where classical controls were not employed (45, 72, 73)
- In a few studies where different control groups were compared, VE estimates were sometimes similar for the test-negative design (74) and sometimes different (75); a comparison of per-protocol results from RCTs versus estimates from test-negative observational studies of an influenza vaccine and of an RSV monoclonal antibody showed close concordance (26) .
- Found that protection appears to decline over the course of influenza season, either due to waning immunity or changing composition of the at-risk group (38-40)
- Complemented other study designs to produce a consistent finding that the 2008-9 seasonal influenza vaccine increased the risk of lab-confirmed clinical infection with 2009 pandemic flu (17)
- Confirmed traditional case-control findings of high VE for monovalent and pentavalent rotavirus vaccines in children (76-79)
- Confirmed traditional case-control findings of likely waning immunity after the 5th dose of pertussis vaccine in children, with similar results to those found with traditional controls (80)
- Suggested that classical case-control design overestimates VE of reduced acellular pertussis (Tdap) vaccines in adolescents and adults (81)
- Confirmed traditional case-control or RCT findings of high VE for pneumococcal vaccines (28, 29, 82)
- Detected low measles VE among children in Democratic Republic of the Congo (83)

Substantive insights from ecologic designs

- Showed that the vaccine-preventable burden of influenza mortality has been overestimated due to confounding (15)

- Showed that conjugate vaccines against Hib and pneumococcal disease have large direct and indirect effects on invasive disease hospitalizations (2, 3, 84)
- Showed reduced benefits due to serotype replacement which led to shift from PCV7 to PCV13(69)
- Demonstrated effect of rotavirus against pediatric diarrhea mortality in Mexico (4)

Table 2. Newer observational study designs and strategies used to measure VE and identify and reduce bias

Strategy	Applicability to types of studies	Summary	Attractive feature	Example of study that used this approach	Conditions for validity	Tradeoffs / Limitations
Negative Control Outcome	Case-control, Cohort, Ecological	Apply parallel analysis to an outcome that should not be affected by vaccine to detect bias in VE estimation	Improves confidence in validity of analytic approach	(19)	Negative control outcome must share confounders of the outcome of interest, but be causally unlinked to vaccine	Can detect direction of bias under some assumptions, but can rarely correct for it
Test-Negative Case-Control	Often seen as a form of case-control study but (26) may be better understood as a form of cohort study (main text)	Choose those who have been tested for clinical syndrome of interest but test negative for target pathogen as controls, to avoid confounding affecting propensity to get tested	Easy, inexpensive access to control group in a surveillance setting	(17, 28, 76-79, 83)	1. vaccine must not affect probability of conditions that lead to testing negative (eg other infections) either directly or indirectly 2. Test highly sensitive and specific	1. Creates selection bias by conditioning on testing, leading to noncausal associations among those tested between vaccination and infection 2. Does not eliminate confounding due to vaccinees having different risk of infection from non-vaccinees, eg. due to place of residence or occupation

Ecological or Trends study:	Ecological study	Compare outcome over time or across populations with varying vaccine coverage to avoid assumption that vaccinated and unvaccinated are comparable	Allows study of total effect of vaccine on population (direct + indirect (herd) effects)	(4)	Any causes of time trends in the outcome other than vaccination must be identified and adjusted for, or extrapolated from time trends before the vaccine is introduced	Major concern is ability to adjust for non-vaccine time trends. Variability in epidemiologic context makes extrapolation of indirect/total effects questionable
------------------------------------	------------------	---	--	-----	--	---