Distinguishing Causation from Correlation in the Use of Correlates of Protection to Evaluate and Develop Influenza Vaccines

Wey Wen Lim, Nancy H. L. Leung, Sheena G. Sullivan, Eric J. Tchetgen Tchetgen, and Benjamin J. Cowling

Correspondence to:

Dr. Nancy H. L. Leung, WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 7 Sassoon Road, Pokfulam, Hong Kong (e-mail: leungnan@hku.hk)

Alternatively,

Prof. Benjamin J. Cowling, WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 7 Sassoon Road, Pokfulam, Hong Kong (e-mail: bcowling@hku.hk)

Author affiliations: WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, The University of Hong Kong, Hong Kong Special Administrative Region, China (Wey Wen Lim, Nancy H. L. Leung and Benjamin J. Cowling); WHO Collaborating Centre for Reference and Research on Influenza at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia (Sheena G. Sullivan); Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, California (Sheena G. Sullivan); Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia (Sheena G. Sullivan); and Statistics Department, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania (Eric J. Tchetgen Tchetgen).

This work received financial support from the Theme-based Research Scheme (grant no. T11-705/14N) from the Hong Kong SAR Government, the National Institute of General Medical Sciences (grant no. U54 GM088558), and a commissioned grant from the Health and Medical Research Fund from the Health, Welfare and Food Bureau of the Hong Kong SAR Government.

Potential conflicts of interest: BJC reports receipt of honoraria from Sanofi and Roche for advisory boards. The authors report no other potential conflicts of interest.

Running head: Causation and Correlation in Vaccinology

Abstract

There is increasing attention on the need to identify new immune markers for the evaluation of existing and new influenza vaccines. Immune markers that could predict individual protection against infection and disease, commonly called "correlates of protection" (CoPs), play an important role in vaccine development and licensing. Here, we discuss the epidemiological considerations when evaluating immune markers as potential CoPs for influenza vaccines and emphasize the distinction between correlation and causation. While an immune marker that correlates well with protection from infection may be used as a predictor of vaccine efficacy, it should be distinguished from an immune marker that plays a mechanistic role in conferring protection against a clinical endpoint, as the latter may be a more reliable predictor of vaccine efficacy and a more appropriate target for rational vaccine design. To clearly distinguish mechanistic and non-mechanistic CoPs, we suggest using the term "correlates of protection" for non-mechanistic CoPs, and "mediators of protection" for mechanistic CoPs. Furthermore, as the interactions among and relative importance of correlates or mediators of protection may vary according to age or prior vaccine experience, the effect sizes and thresholds for protective effects for CoPs may also vary in different segments of the population.

Keywords: Influenza, Human; Influenza Vaccines; Biomarkers; Terminology as Topic; Causality; immune correlates of protection; immune mediators of protection; immune markers

Abbreviations:

CoP correlate of protection

HAI haemagglutination inhibition

MoP mediator of protection

LAIV live attenuated influenza vaccine

Introduction

Immune markers that could predict individual protection against infection and disease, commonly called "correlates of protection" (CoPs), play an important role in vaccine development and licensing (1). As measures of immunogenicity, CoPs can provide a useful indication of protective biological responses induced by investigational vaccines before they are tested in large efficacy trials (2). As substitute endpoints for clinical outcomes, CoPs can expedite the licensure of vaccines based on their association with increased levels of established CoPs. Pandemic influenza vaccines, for instance, can be licensed based on their association with increased levels of immune biomarkers (immune CoPs) without requiring a demonstration of actual protection against clinical outcomes (3,4).

The importance of appropriate and validated CoPs in vaccine development is well recognized. Yet, their identification and validation remain difficult for many diseases, including tuberculosis, malaria, and human immunodeficiency virus infections (2,5-9). For influenza, there is an increasing body of research dedicated to understanding the immune response to influenza exposures aimed at identifying new immune markers for the evaluation of existing and new influenza vaccines (9-11). These include systems biology studies of post-vaccination immune profiles (12,13), and experimental or observational studies to correlate various immune markers with clinical endpoints such as influenza virus infection, disease, and virus shedding (14-18).

Nevertheless, there is a need to establish the causal contribution of these immune markers to protection against clinical endpoints because interventions targeted at non-causal predictors may be ineffective. Here, our objective is to review the analytical frameworks used to establish an

immune marker as a CoP, advocate for the use of the term "mediator of protection" (MoP), and discuss methods and important considerations when evaluating an immune marker as a potential CoP or MoP to assess immunity against influenza virus infections and disease.

Correlates of protection (CoPs) play an important role in influenza vaccine evaluation

Historically, the development and use of CoPs as substitutes or surrogates for "true" clinical endpoints occur when the clinical endpoints are rare, require a long time to develop, or are expensive or difficult to measure (19). For infectious diseases, CoPs are important when post-intervention outcomes of interest can be difficult to detect, for example either asymptomatic or transient infections (e.g. influenza virus infections, latent tuberculosis), or have long incubation periods (e.g. acquired immunodeficiency syndrome after human immunodeficiency virus infections). CoPs have been used as indicators of protective immunity and immune responses for example those generated by vaccination, and can be measured to evaluate the effectiveness of existing prevention or treatment strategies, and as targets to develop new interventions (20,21).

Influenza vaccines can be licensed based only on clinical efficacy, and without a corresponding CoP. One example of this is the live attenuated influenza vaccine (LAIV) FluMist (MedImmune Inc., Gaithersburg, MD), which was licensed in the United States in 2003 based on its clinical efficacy in large randomized placebo-controlled trials (22). However, CoPs are important for the evaluation of current inactivated influenza vaccines, as they provide a relatively rapid and cost-effective measure of immune responses to seasonal and pandemic influenza vaccines, circumventing the need for large efficacy trials when seasonal vaccines are updated annually and when pandemic vaccines need to be evaluated and licensed quickly before deployment. Three

CoPs are established and used by regulatory agencies for licensure of inactivated influenza vaccines. These are the antibody titers measured by the hemagglutination inhibition (HAI) assay, the single radial hemolysis assay, and the virus neutralization or microneutralization assay (6).

The search for additional CoPs for influenza vaccines in humans

In recent years, the sufficiency of current CoPs has been called into question (6,9). In the case of HAI titers, failures of inactivated influenza vaccination to prevent laboratory-confirmed infection despite high post-vaccination HAI titers against influenza viruses included in the vaccine have been documented, particularly for A(H3N2) viruses (23). While the microneutralization assay provides a direct measure of antibodies that inhibit influenza viral entry, the assay is labor intensive, can have poor inter-laboratory reproducibility (6,24), and currently has no established threshold for protection (24). Notwithstanding the technical challenges and limitations of these CoPs, the highly mutable influenza virus also poses an additional challenge to the use of post-vaccination CoPs to evaluate influenza vaccines. As circulating influenza virus strains can easily mutate and acquire antigenic changes that can escape host immune responses, post-vaccination antibody titers such as HAI titers that are induced specifically against vaccine strains can be less effective in preventing infections by circulating strains that may have evolved significantly from the vaccine strains (9,25). This is in contrast to vaccinations against infections such as measles, of which exposures, if overcome, will induce immune responses that provide long-lasting immunity (25). Moreover, next-generation universal influenza vaccines aiming to prove broader protection are likely to work through other immune mechanisms, and thus would not be properly evaluated with HAI-based assays. These underscore the need for additional markers that can

allow a more comprehensive evaluation of protection conferred by current and next-generation influenza vaccines.

Additional candidates of CoPs for current and next-generation influenza vaccines have been suggested elsewhere, including antibodies against the stalk of the influenza hemagglutinin protein, antibodies against other influenza virus surface proteins, components of cell-mediated immunity, and components of mucosal immunity (18,24). Several clinical studies have measured these alternative immune markers after influenza virus infection or vaccination, assessed their relationships with subsequent infections, viral shedding, symptom duration and scores, hospitalization, or death, and reported on factors that may affect these relationships (14,18,24,26). These factors include the route of vaccine administration, mode of infection (experimental vs. natural), antigenic match between vaccine and circulating virus strains, participant characteristics (e.g. age), immune status (e.g. pre-existing immunity), and sample size (14,27). However, few studies have attempted to quantitatively evaluate the strength or validity of these alternative immune measures as CoPs (28,29) or their relative importance in protection.

The evolution of statistical approaches for evaluating CoPs

The practical benefits of using CoPs as substitute endpoints have facilitated their use in clinical trials. However, statistical methods for identifying CoPs can inadequately predict the intervention's effect on true endpoints such as survival or disease progression (19). This can be problematic if an intervention is designed to modify levels of the CoP under the erroneous assumption that it will have a causal effect on a person's risk of developing the disease. An example can be found in tuberculosis vaccine research, when researchers did not observe

significant protective efficacy against tuberculosis infection in infants receiving a candidate booster vaccine in a phase 2b trial (30), after the vaccine have elicited significantly higher levels of an immune marker (31) that is thought to be a strong CoP against tuberculosis infections in a phase I clinical trial (5,31).

To overcome the limitations of relying on statistical associations to establish the validity of CoPs as substitute endpoints, Prentice introduced a set of criteria that should be met for a CoP to be called a "surrogate" or a substitute endpoint to a "true" endpoint of interest (**Figure 1**) (32). Prentice proposed that for a CoP to be a valid surrogate for an effective intervention, its levels must not only be significantly related to the intervention and the true endpoint; it must "capture any relationship between the treatment and the true endpoint" in the context of randomized controlled trials (**Figure 1A**) (32). Therefore, a surrogate is a CoP that will nullify any association between intervention and a clinical endpoint when it is accounted for in a statistical model (**Figure 1B**) (32). This criteria is quite restrictive, as few interventions work through a single mechanism or causal pathway (1,33).

Another approach proposed in the 1990s is based on the "proportion of treatment effect explained" (PTE) by a CoP (34,35). Here, the expectation is that a valid surrogate endpoint can be a CoP that "accounts for a substantial portion of treatment effect on the clinical endpoint" (36), even when the surrogacy is "incomplete" (37). Some of the limitations of this approach, detailed in De Gruttola et al.'s summary of a National Institutes of Health Workshop on surrogate endpoints, include the difficulty in interpreting the proportion of treatment effect explained (38). For instance, the relative magnitude of the treatment effect on the true and surrogate endpoint,

and the association of the surrogate and true endpoint independent of treatment are important in determining the utility and validity of the surrogate endpoint, but this is usually not considered in methods that estimate the proportion of treatment effect explained (39). High values of the proportion of treatment effect explained also do not imply that a surrogate endpoint lies on a causal pathway from an intervention to a clinical endpoint, unless the relationship is modelled perfectly (38).

In 2002, Frangakis and Rubin proposed the idea of a "principal surrogate" to assess the validity of a CoP based on its causal nature (40,41). This is a CoP for which the effect of treatment on the true endpoint is the same within categories of principal strata (categories based on variables that are not affected by treatment assignment) for fixed levels of a CoP, thereby demonstrating the "causal necessity" property of the CoP (40,42). An example of this would be when the same value of the CoP corresponds to the same strength of protection for people of different ages or vaccination histories. These properties are assessed by characterizing changes in treatment efficacy with subgroup analyses on groups categorized into principal strata, which led Gilbert et al. to suggest "principal stratification effect modification analysis" as a name to describe such analysis (42). A principal surrogate is a CoP in which modification of the relationship between an intervention and outcome is consistent among pre-treatment variables. However, the principal surrogate property may not be met by some causal CoPs (43), and may be too restrictive in scenarios where multiple mechanisms operate. In some scenarios, an intervention may have a negative impact on the clinical endpoint despite demonstrating a positive effect or is positively associated with the surrogate endpoint (44).

Terminology for correlates of protection

These developments in the concept of surrogacy have been accompanied by confusion in the terminology used to describe CoPs. While the National Institutes of Health Biomarkers Definitions Working Group had proposed definitions for the terms "biological marker" (biomarker), "clinical endpoint", and "surrogate endpoint" in 2001 (45), both "correlates" and "surrogates" are used to describe substitute endpoints in scientific and regulatory documents, and other terms such as "immune marker of protection" and "intermediate endpoint" are still commonly used (1). In 2007, Qin et al. proposed a framework in which a potential "correlate of risk" (CoR), or an immunological measurement that could predict a clinical endpoint, could be assessed. They proposed 2 categories, "level 1 surrogate of protection" and "level 2 surrogate of protection", based on whether correlations between levels of the correlate of risk and clinical endpoints can be found in both the vaccinated and unvaccinated (level 1), and whether this correlation can be replicated in different populations or settings (level 2), after satisfying the Prentice criteria (46,47). However, they acknowledged that the identification and validation of an immune CoP that causes or mediates vaccine-induced protection likely requires a mechanistic understanding, and CoPs that only partially contribute to a vaccine's protective effect will not be identified under this paradigm (46). Plotkin and Gilbert highlighted the need to distinguish mechanistic CoPs (mCoPs), from non-mechanistic CoPs (nCoPs). The latter do not play a causal role in protection against a specific clinical endpoint but predict protection through correlations with other causal protective immune responses (48).

While the use of the terms mechanistic CoPs and non-mechanistic CoPs are useful in differentiating CoPs that are likely causal from the non-causal, as described by Plotkin and

Gilbert (48), we believe there is a simpler and more direct way to describe a biomarker that plays a causal role in protection against a specific clinical endpoint. Since it is generally of most interest to determine whether an immune marker merely predicts or causes protection against a clinical endpoint, a non-mechanistic CoP could still be called a "correlate of protection" (CoP), as it describes exclusively a *statistical association* between vaccination and risk of infection or disease (48). A mechanistic CoP could be called a "mediator of protection" (MoP) to better reflect its role on a *causal pathway* between the two.

The difference between a CoP and MoP can be illustrated with a directed acyclic graph (DAG; see Web Appendix for basic notation and terminology), which is a type of causal diagram and is a graphical visualization of the assumed relationships between exposures, outcomes, and other factors related to exposure and/or outcome (**Figure 2**). According to our definitions, a CoP may or may not lie on a direct causal pathway between natural infection or vaccination and protection against a clinical endpoint, while a MoP must lie on such a pathway (**Figure 2**).

Identifying immune MoPs for influenza vaccines using causal inference frameworks

While many potential CoPs for influenza vaccines have been proposed, we believe that the focus should shift to the identification of new MoPs for evaluation of next-generation influenza vaccines, especially when the ultimate goal of these vaccines is to generate broader and longer-lasting protection (9). While it is a reasonable approach to design vaccines that aim to generate immune response against conserved viral targets, we would be remiss if we do not also consider the biological plausibility or capacity of the consequent immune response to generate a protective effect against infection or disease. This could potentially be done through laboratory

investigation into the mechanisms of protective immune responses, and further assessment of the causal role of these CoPs in human studies through causal inference frameworks.

Compared with standard statistical methods to assess the association between a CoP and a clinical endpoint, causal analysis approaches such as counterfactual-based mediation analysis methods have the advantage of accommodating formal hypotheses (49-51) about the causal contribution of candidate MoPs to the risk of infection. These hypotheses are based on current understanding of immune mechanisms and can be tested quantitatively. A CoP for an influenza vaccine that is also on the causal pathway for protection against infection or disease can be described as a MoP, which is a component of the immune response, which if not present or stimulated after vaccination, will result in elimination (as the protection is fully mediated by the MoP) (**Figure 3A** and **Figure 3B**) or attenuation (as the protection is partially mediated by the MoP) of vaccine-induced protection (**Figure 3C** and **Figure 3D**).

In causal mediation analyses where the effect of vaccination on a clinical endpoint is partially mediated by a particular immune marker (**Figure 3C**), we can decompose the "total effect" of vaccination on the clinical endpoint into at least two components – the indirect and direct effects. The "indirect effect" of vaccination (V) on the clinical endpoint (I) is the effect that is mediated by the immune marker (M) under study (the path V → M → I in **Figure 3C**). Conversely, the "direct effect" of vaccination on a clinical endpoint is the effect that is not mediated by the immune marker under study (the path consists of a single edge V → I in **Figure 3C**). In the simplest possible of settings, Baron and Kenny's approach (52) can sometimes be used to perform causal mediation analysis, under fairly strong assumptions of no effect modification

between the mediator and the exposure in the outcome model and stringent no unmeasured confounding assumption. This approach entails estimating the direct effect by including both the exposure and the mediator in a regression model for the outcome, and estimating the indirect effect by subtracting the direct effect from the total effect obtained by removing the mediator from the regression. In recent years, there has been a growing recognition that the Baron and Kenny approach is often not appropriate, and a more general counterfactual framework for mediation analysis has been adopted (50,53-55). Within a counterfactual framework, the causal effect of an intervention is conceptualized as the difference between two "potential outcomes" or "counterfactual outcomes" (56). While these two outcomes by definition cannot be observed simultaneously for the same individual, the average causal effect for a specific study population can be estimated by comparing these counterfactual outcomes for that study population (57). This is in contrast to the use of separate regression models that relate exposure to a CoP, and relate CoP to a clinical outcome to identify CoPs. Under this causal mediation framework we can formally recognize the assumptions that are essential for the estimation of direct and indirect effects (also known as natural direct and indirect effects), including the assumption of no-unmeasured confounding of the relationships between exposure and outcome (V and I), exposure and mediator (V and M), and between mediator and outcome (M and I) (58). The assumption of no unmeasured confounding between mediator and outcome may be violated if an unmeasured immune mechanism induced by the exposure influences levels of the immune marker under investigation as well as the clinical endpoint. Although, as recently shown by Fulcher et al. (59), progress can sometimes be made by an appropriate analytic method even when such unmeasured confounding exists.

A more specific example would be evaluating potential MoPs for LAIV (**Figure 4)**. Although the serum HAI antibody was once expected to be a strong MoP for LAIV (**Figure 4A**), as believed to be the case for inactivated influenza vaccines, a study found that the serum HAI antibody titer underestimates protection (60), and another study suggested that mucosal secretory IgA (sIgA) antibody may be a more appropriate MoP (61). Based on the current understanding of the mechanisms of immune responses generated by LAIV (61), there may be several hypotheses for this observation. For example, LAIV may confer protection through both mucosal secretory IgA antibody and, to a lesser extent, serum HAI antibody (**Figure 4B**). Alternatively, it may confer protection only through mucosal secretory IgA and other immune mechanisms that do not affect the levels of serum HAI antibody (**Figure 4C**). One could test these hypotheses by drawing DAGs to describe the hypothesized causal relationships between factors and formulating statistically testable hypotheses based on these relationships.

In a recent study, we demonstrated the use of a causal analytical approach to quantitatively evaluate the causal contribution of HAI antibody titers to protection against influenza B virus infection, in a randomized placebo-controlled trial of inactivated influenza vaccines in children 6-17 years old (62). In the study, using inverse odds ratio weighting, we estimated that the post-vaccination HAI titer mediated 57% of the causal effect of inactivated influenza vaccines on protection against influenza B virus infection. However, this study utilized just one immune measure, which is the serum HAI antibody titer measured by the HAI assay, and just one reference antigen used in the HAI assay. The remaining 43% could be explained by residual HA-targeting antibody response not captured in that HAI assay, or other components of the immune response.

As more potential CoPs and MoPs are identified and evaluated, it is also important to note that the role as a CoP or MoP of an immune measure may differ between the way it is generated, e.g. whether it is by influenza virus infections or vaccinations. These roles may also differ according to the endpoints being measured (63), type of vaccines (64) and population characteristics (65). As the interactions among and relative importance of CoPs and MoPs may vary according to historical exposures to influenza including prior vaccinations or infections, the effect sizes and thresholds required for the strength of association for CoPs or the protective effect for MoPs may also vary in different segments of the population. As such, established and novel CoPs and MoPs need to be identified and evaluated in different population groups, such as in different age groups, as well as for different influenza strains.

Conclusion

In conclusion, identifying new correlates and mediators of immune protection is a critical step for the development and evaluation of next-generation and universal influenza vaccines. It will be important to collect data on multiple immune measures from the same study and decipher their relative causal contribution to protection. Although we may still be unable to isolate the causal contribution of a single immune marker to protection with such studies, we may be able to indicate that a composite MoP consists of several immune markers may be a better MoP than any single immune marker. Given the value of causal immune markers in vaccine evaluation, more research is needed to identify CoPs which mediate protection for next-generation and universal influenza vaccines. A unifying term to describe these immune markers, such as the term "mediator of protection" suggested here, can be an important first step to raise awareness of the

need for causal evaluation of CoPs and stimulate discussion on the desired characteristics of MoPs for future influenza vaccines. The eventual adoption of any immune marker or group of markers as a tool to evaluate new influenza vaccines would have to consider both their reliability as predictors of vaccine efficacy and the cost and technical demands of their measurement. Although there is also not thought to be any natural long-lasting broadly cross-reactive immunity against influenza in humans, as indicated by repeated influenza virus infections during a person's lifetime, one may question if it will be possible to identify a MoP for universal influenza vaccines. As repeated influenza virus infections could be due to antigenic drift and the increasing mismatch between circulating viruses and host immunity (i.e. reduced cross-protection/ a limitation on the breadth of immune responses) and/or a decreasing level of immunity over time (i.e. a limitation on the duration of immune responses), an ideal MoP for universal influenza vaccines would need to first be shown to be cross-reactive, and subsequently a vaccine designed based on this MoP to demonstrate a persistent level of the MoP that may be maintained after a single or repeated doses. A MoP that is only cross-reactive may not be ideal, but would still be very useful for novel pandemics where the concern is mainly on the very limited population immunity against the novel virus.

Acknowledgements

We thank Julie Au for administrative support.

References

1. World Health Organization. Correlates of vaccine-induced protection: methods and implications. 2013. https://www.who.int/immunization/documents/WHO_IVB_13.01/en/. Accessed February 20, 2019.
2. van Els C, Mjaaland S, Næss L, et al. Fast vaccine design and development based on correlates of protection (COPs). *Hum Vaccin Immunother.* 2014;10(7):1935-1948.
3. Greenberg ME, Lai MH, Hartel GF, et al. Response to a monovalent 2009 influenza A (H1N1) vaccine. *N Engl J Med.* 2009;361(25):2405-2413.
4. Weir JP, Gruber MF. An overview of the regulation of influenza vaccines in the United States. *Influenza Other Respir Viruses.* 2016;10(5):354-360.
5. Bhatt K, Verma S, Ellner JJ, Salgame P. Quest for Correlates of Protection against Tuberculosis. *Clin Vaccine Immunol.* 2015;22(3):258-266.
6. Trombetta CM, Montomoli E. Influenza immunology evaluation and correlates of protection: a focus on vaccines. *Expert Rev Vaccines.* 2016;15(8):967-976.
7. Tomaras GD, Plotkin SA. Complex immune correlates of protection in HIV-1 vaccine efficacy trials. *Immunol Rev.* 2017;275(1):245-261.
8. Laurens MB. The Immunologic Complexity of Growing Up with Malaria—Is Scientific Understanding Coming of Age? *Clin Vaccine Immunol.* 2016;23(2):80.
9. Erbelding EJ, Post DJ, Stemmy EJ, et al. A universal influenza vaccine: The strategic plan for the National Institute of Allergy and Infectious Diseases. *J Infect Dis.* 2018:347-354.
10. Mohn KGI, Smith I, Sjursen H, Cox RJ. Immune responses after live attenuated influenza vaccination. *Hum Vaccin Immunother.* 2018;14(3):571-578.

11. Krammer F. The human antibody response to influenza A virus infection and vaccination. *Nat Rev Immunol.* 2019;19(6):383-397.

12. Nakaya HI, Wrammert J, Lee EK, et al. Systems biology of vaccination for seasonal influenza in humans. *Nat Immunol.* 2011;12(8):786-795.

13. Tsang JS, Schwartzberg PL, Kotliarov Y, et al. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell.* 2014;157(2):499-513.

14. Valkenburg SA, Leung NHL, Bull MB, et al. The hurdles from bench to bedside in the realization and implementation of a universal influenza vaccine. *Front Immunol.* 2018;9(1479).

15. Clements ML, Betts RF, Tierney EL, Murphy BR. Serum and nasal wash antibodies associated with resistance to experimental challenge with influenza A wild-type virus. *J Clin Microbiol.* 1986;24(1):157-160.

16. Hayward A, Wang L, Goonetilleke N, et al. Natural T cell-mediated protection against seasonal and pandemic Influenza. *Am J Respir Crit Care Med.* 2015;191(12):1422-1431.

17. Wilkinson TM, Li CK, Chui CS, et al. Preexisting influenza-specific CD4+ T cells correlate with disease protection against influenza challenge in humans. *Nat Med.* 2012;18(2):274-280.

18. Li CK-f, Rappuoli R, Xu X-N. Correlates of protection against influenza infection in humans—on the path to a universal vaccine? *Curr Opin Immunol.* 2013;25(4):470-476.

19. Molenberghs G, Buyse M, Burzykowski T. *The evaluation of surrogate endpoints.* New York, NY: Springer New York; 2005.

20. Plotkin SA. Vaccines: correlates of vaccine-induced immunity. *Clin Infect Dis.* 2008;47(3):401-409.

21. Katz JM, Hancock K, Xu X. Serologic assays for influenza surveillance, diagnosis and vaccine evaluation. *Expert Rev Anti Infect Ther.* 2011;9(6):669-683.

22. Harper SA, Fukuda K, Cox NJ, Bridges CB. Using live, attenuated influenza vaccine for prevention and control of influenza: supplemental recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Recomm Rep.* 2003;52(RR-13):1-8.

23. Ohmit SE, Petrie JG, Cross RT, Johnson E, Monto AS. Influenza hemagglutination-inhibition antibody titer as a correlate of vaccine-induced protection. *J Infect Dis.* 2011;204(12):1879-1885.

24. Gianchecchi E, Torelli A, Montomoli E. The use of cell-mediated immunity for the evaluation of influenza vaccines: an upcoming necessity. *Hum Vaccin Immunother.* 2019;15(5):1021-1030.

25. Bouvier NM. The Future of Influenza Vaccines: A Historical and Clinical Perspective. *Vaccines.* 2018;6(3):58.

26. Boyce TG, Gruber WC, Coleman-Dockery SD, et al. Mucosal immune response to trivalent live attenuated intranasal influenza vaccine in children. *Vaccine.* 1999;18(1-2):82-88.

27. Tsang JS. Utilizing population variation, vaccination, and systems biology to study human immunology. *Trends Immunol.* 2015;36(8):479-493.

28. Monto AS, Petrie JG, Cross RT, et al. Antibody to influenza virus neuraminidase: an independent correlate of protection. *J Infect Dis.* 2015;212(8):1191-1199.

29. Park J-K, Han A, Czajkowski L, et al. Evaluation of preexisting anti-hemagglutinin stalk antibody as a correlate of protection in a healthy volunteer challenge with influenza A/H1N1pdm virus. *mBio.* 2018;9(1).

30. Tameris MD, Hatherill M, Landry BS, et al. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. *Lancet.* 2013;381(9871):1021-1028.

31. McShane H, Pathan AA, Sander CR, et al. Recombinant modified vaccinia virus Ankara expressing antigen 85A boosts BCG-primed and naturally acquired antimycobacterial immunity in humans. *Nat Med.* 2004;10(11):1240-1244.

32. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med.* 1989;8(4):431-440.

33. De Gruttola V, Fleming T, Lin DY, Coombs R. Perspective: Validating surrogate markers: Are we being naive? *J Infect Dis.* 1997;175(2):237-246.

34. Freedman LS, Graubard BI, Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. *Stat Med.* 1992;11(2):167-178.

35. Schatzkin A, Freedman LS, Schiffman MH, Dawsey SM. Validation of intermediate end points in cancer research. *J Natl Cancer Inst.* 1990;82(22):1746-1752.

36. Lin DY, Fleming TR, De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. *Stat Med.* 1997;16(13):1515-1527.

37. Choi S, Lagakos SW, Schooley RT, Volberding PA. CD4+ lymphocytes are an incomplete surrogate marker for clinical progression in persons with asymptomatic HIV infection taking zidovudine. *Ann Intern Med.* 1993;118(9):674-680.

38. De Gruttola VG, Clax P, DeMets DL, et al. Considerations in the Evaluation of Surrogate Endpoints in Clinical Trials: Summary of a National Institutes of Health Workshop. *Control Clin Trials.* 2001;22(5):485-502.

39. Buyse M, Molenberghs G. Criteria for the validation of surrogate endpoints in randomized experiments. *Biometrics.* 1998;54(3):1014-1029.

40. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics.* 2002;58(1):21-29.

41. Gilbert PB, Hudgens MG. Evaluating candidate principal surrogate endpoints. *Biometrics.* 2008;64(4):1146-1154.

42. Gilbert PB, Gabriel EE, Huang Y, Chan IS. Surrogate Endpoint Evaluation: Principal Stratification Criteria and the Prentice Definition. *J Causal Inference.* 2015;3(2):157-175.

43. Black S, Nicolay U, Vesikari T, et al. Hemagglutination inhibition antibody titers as a correlate of protection for inactivated influenza vaccines in children. *Pediatr Infect Dis J.* 2011;30(12):1081-1085.

44. Vanderweele TJ. Surrogate measures and consistent surrogates. *Biometrics.* 2013;69(3):561-569.

45. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001;69(3):89-95.

46. Qin L, Gilbert PB, Corey L, McElrath MJ, Self SG. A framework for assessing immunological correlates of protection in vaccine trials. *J Infect Dis.* 2007;196(9):1304-1312.

47. Gilbert PB, Qin L, Self SG. Evaluating a surrogate endpoint at three levels, with application to vaccine development. *Stat Med.* 2008;27(23):4758-4778.

48. Plotkin SA, Gilbert PB. Nomenclature for immune correlates of protection after vaccination. *Clin Infect Dis.* 2012;54(11):1615-1617.

49. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology.* 1992;3(2):143-155.

50. Pearl J. Direct and Indirect Effects (UA1-01). Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence; 2001; San Francisco, CA.

51. Avin C, Shpitser I, Pearl J. Identifiability of path-specific effects. Proceedings of the 19th International Joint Conference on Artificial Intelligence; July 30-August, 2005; Edinburgh; Scotland, UK. .

52. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol.* 1986;51(6):1173-1182.

53. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology.* 1999;10(1):37-48.

54. Pearl J. The foundations of causal inference. *Soc Method.* 2010;40(1):75-149.

55. VanderWeele TJ. *Explanation in causal inference : methods for mediation and interaction.* New York, NY : Oxford University Press; 2015.

56. Robins J, Hernán MA. *Causal inference, Part I.* Boca Raton: Chapman & Hall/CRC; [available online ahead of print, September 7, 2019 at https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1268/2019/09/hernanrobins_v1.10.42.pdf].

57. VanderWeele TJ, Vansteelandt SWE. Conceptual issues concerning mediation, interventions and composition. *Stat Interface.* 2009;2:457-468.

58. VanderWeele TJ. Mediation Analysis: A Practitioner's Guide. *Annu Rev Public Health.* 2016;37(1):17-32.

59. Fulcher IR, Shi X, Tchetgen Tchetgen EJ. Estimation of natural indirect effects robust to unmeasured confounding and mediator measurement error. *Epidemiology.* [available online ahead of print August 30, 2019 at doi: 10.1097/EDE.0000000000001084].

60. Cox RJ. Correlates of protection to influenza virus, where do we go from here? *Hum Vaccin Immunother.* 2013;9(2):405-408.

61. Jin H, Subbarao K. Live attenuated influenza vaccine. *Curr Top Microbiol Immunol.* 2015;386:181-204.

62. Cowling BJ, Lim WW, Perera RAPM, et al. Influenza hemagglutination-inhibition antibody titer as a mediator of vaccine-induced protection for influenza B. *Clin Infect Dis.* 2018;68(10):1713-1717.

63. Memoli MJ, Shaw PA, Han A, et al. Evaluation of antihemagglutinin and antineuraminidase antibodies as correlates of protection in an influenza A/H1N1 virus healthy human challenge model. *mBio.* 2016;7(2):e00417-00416.

64. Hoft DF, Babusis E, Worku S, et al. Live and inactivated influenza vaccines induce similar humoral responses, but only live vaccines induce diverse T-cell responses in young children. *J Infect Dis.* 2011;204(6):845-853.

65. Dunning AJ, DiazGranados CA, Voloshen T, Hu B, Landolfi VA, Talbot HK. Correlates of protection against influenza in the elderly: results from an influenza vaccine efficacy trial. *Clin Vaccine Immunol.* 2016;23(3):228-235.

Figure Legends

Figure 1. A causal diagram depicting the relationship between an intervention, a surrogate, and a clinical endpoint according to the Prentice criteria in the context of an influenza vaccine randomized-controlled trial (32). According to Prentice, a surrogate for an effective intervention needs to satisfy four criteria: (1) the intervention (vaccination) must be significantly associated with the clinical endpoint (infection); (2) the intervention (vaccination) must be significantly associated with the surrogate; (3) the surrogate must be significantly associated with the clinical endpoint (infection); (4) the clinical endpoint (infection) is independent of the intervention (vaccination) conditional on the surrogate variable. In **Panel A**, influenza vaccination is expected to have an effect on the surrogate ("a"), which in turn is expected to have an effect on influenza virus infection ("b"). Therefore, influenza vaccination is expected to have an effect on influenza virus infection through the surrogate marker (a + b), satisfying criteria 1-3. In **Panel B**, as influenza vaccination is expected to have an effect on influenza virus infection only through the surrogate, influenza vaccination and influenza virus infection would be independent of each other if the surrogate variable is accounted for in a statistical model. By accounting for the surrogate variable, the surrogate variable is prevented from varying by vaccination status. Hence, the arrow from vaccination to surrogate is removed.

Figure 2. A causal diagram representing the hypothesis regarding the causal relationships between historical exposures to influenza (either by historical infections, or historical vaccinations with inactivated or attenuated virus) (H), current influenza vaccination (V), a correlate of protection (CoP), a mediator of protection (MoP), and a clinical endpoint such as infection, disease, or influenza-related mortality (I). In causal framework terminology, here in

the evaluation of the protective effect of vaccination (V) on clinical outcomes/ endpoints such as infection (I), V takes the role of cause, I as effect, MoP as mediator, and H as confounder (common cause) to MoP and I. The *path* consists of a single-headed arrow ("edge") drawn from V to I represents a "direct effect" of vaccination on infection (V → I), while "indirect effect" of V on I is represented by the path that consists of the two edges from V to MoP and from MoP to I (V → MoP → I), with MoP being the mediator of such effect. The "total effect" of V on I is the sum of the direct and indirect effect represented by the two paths. An absence of a *directed path* between two factors represent an assumption of no causal relationship between the two, for example between CoP and I. While a CoP can lie on both causal and non-causal pathways, a MoP is a CoP that lies on a causal pathway between vaccination (or historical exposures to influenza) and infection. An association between the CoP and (reduced) risk of infection can still be observed due to confounding by historical exposures to influenza or current influenza vaccination, or if another unobserved immune marker (as represented by the *absence* of it in the diagram) affects both the level of CoP and the risk of infection. A brief introduction including graphical notation and terminology of directed acyclic graphs (DAGs), a type of causal diagram, are available in **Web Appendix**.

Figure 3. Causal diagrams depicting two hypotheses of the role of an immune marker (M) as a mediator of protection (MoP) against clinical endpoints such as infection (I), as influenced by the effect of current influenza vaccination (V) or historical exposures to influenza infection/vaccination (H). **Panels A** and **B** depicts **full mediation** by immune marker M on protection against infection I before and after adjustment for M respectively. In **Panel A**, the protective effect of historical exposures to influenza or current influenza vaccination (H and V)

on I is mediated only by M, indicated by the two paths H → M → I and V → M → I respectively. As such, as depicted in **Panel B**, the protective effect of historical exposures or current vaccination will not be observed, in theory, after M is controlled or adjusted for in a statistical model as graphically represented by a box around M, i.e. M can explain all the effect of H or V on I. By controlling or adjusting for M, values of M are prevented from varying by values of V or H. Therefore, the edges V → M and H → M are removed. **Panels C** and **D** depicts **partial mediation** by immune marker M on protection against infection I before and after adjustment for M respectively. In **Panel C**, the protective effect of historical exposures to influenza or current influenza vaccination (H and V) is mediated by M as well as other unmeasured immune responses (as unobserved confounders in the causal context), with the effect of historical exposures to influenza indicated by the two paths H → M → I (i.e. mediated by M) and H → I (i.e. mediated by unmeasured immune responses), and the effect of current influenza vaccination by the other two paths V → M → I and V → I. In this context, as depicted in **Panel D**, the protective effect of historical exposures to influenza or current influenza vaccination will still be observed even when M is controlled or adjusted for in a statistical model, through the protection mediated by the unmeasured immune responses as represented by the two paths H → I and V → I respectively, i.e. M can only partially explain all the effect of H or V on I.

Figure 4. Causal diagrams of three hypotheses of the causal relationships between vaccination with live attenuated influenza vaccines (LAIV), a clinical endpoint such as infection (I), and two potential mediators of protection (MoPs) hemagglutination inhibition antibody titer (HAI) and secretory IgA antibodies (sIgA). In the first hypothesis as depicted in **Panel A**, live attenuated influenza vaccine (LAIV) is hypothesized to confer protection though HAI antibody titer and

other immune mechanisms. Therefore, there is a path LAIV → HAI → I, indicating an indirect effect through HAI, and a path LAIV → I, indicating the protective effect from other non-HAI-mediated immune mechanisms. In the second hypothesis as depicted in **Panel B**, LAIV is hypothesized to confer protection through sIgA and HAI as well as other immune mechanisms, hence there are three paths from LAIV to I, including the path mediated through sIgA (LAIV → sIgA → I), the path mediated through HAI (LAIV → HAI → I), as well as the path consists of a single edge LAIV → I. In the third hypothesis as depicted in **Panel C**, LAIV confers protection only through sIgA and other non-HAI-mediated immune mechanisms, hence there are two paths from LAIV to I only (LAIV → sIgA → I and LAIV → I), while the path LAIV → HAI → I is absent due to the absence of the edge LAIV → HAI.