



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Conference report

Evaluation of post-introduction COVID-19 vaccine effectiveness: Summary of interim guidance of the World Health Organization

Minal K. Patel^{a,*}, Isabel Bergeri^a, Joseph S. Bresee^b, Benjamin J. Cowling^c, Natasha S. Crowcroft^a, Kamal Fahmy^d, Siddhivinayak Hirve^a, Gagandeep Kang^e, Mark A. Katz^f, Claudio F. Lanata^g, Maïna L'Azou Jackson^h, Sudhir Joshiⁱ, Marc Lipsitch^j, Jason M. Mwenda^k, Francisco Nogareda^l, Walter A. Orenstein^m, Justin R. Ortizⁿ, Richard Pebody^f, Stephanie J. Schrag^b, Peter G. Smith^o, Padmini Srikantiah^p, Lorenzo Subissi^a, Marta Valenciano^q, David W. Vaughn^p, Jennifer R. Verani^b, Annelies Wilder-Smith^a, Daniel R. Feikin^a

^a World Health Organization, 20 Avenue Appia, Geneva 1211, Switzerland

^b U.S. Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA, USA

^c School of Public Health, The University of Hong Kong, Pokfulam, Hong Kong, China

^d World Health Organization Regional Office for the Eastern Mediterranean, Monazamet El Seha El Alamia Str, Extension of Abdel Razak El Sanhoury Street, P.O. Box 7608, Nasr City, Cairo 11371, Egypt

^e Christian Medical College, Ida Scudder Road, Vellore, Tamil Nadu 632004, India

^f World Health Organization Regional Office of Europe, UN City, Marmorvej 51, Copenhagen DK-2100, Denmark

^g Instituto de Investigación Nutricional, Av. la Molina 1885, La Molina 15024, Peru

^h The Coalition for Epidemic Preparedness Innovations (CEPI), Gibbs building, 215 Euston Rd, Bloomsbury, London NW1 2BE, United Kingdom

ⁱ World Health Organization Regional Office for South-East Asia, World Health House, Indraprastha Estate, Mahatma Gandhi Marg, New Delhi 110 002, India

^j Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA

^k World Health Organization Regional Office for Africa, Cité du Djoué, P.O. Box 06, Brazzaville, Republic of Congo

^l Consultant to the Pan American Health Organization, 525 23rd Street NW, Washington, DC 20037, USA

^m Emory Vaccine Center, 954 Gatewood Road, Atlanta, GA 30329, USA

ⁿ Center for Vaccine Development & Global Health, University of Maryland School of Medicine, 685 W. Baltimore St., Room #480, Baltimore, MD 21201, USA

^o MRC International Epidemiology & Statistics Group, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom

^p Bill & Melinda Gates Foundation, 500 5th Ave N., Seattle, WA 98109, USA

^q Epiconcept, 25 rue de Titon, Paris 75011, France

ARTICLE INFO

Article history:

Received 1 April 2021

Accepted 27 May 2021

Available online 1 June 2021

Keywords:

COVID-19

Vaccination

Vaccine effectiveness

ABSTRACT

Phase 3 randomized-controlled trials have provided promising results of COVID-19 vaccine efficacy, ranging from 50 to 95% against symptomatic disease as the primary endpoints, resulting in emergency use authorization/listing for several vaccines. However, given the short duration of follow-up during the clinical trials, strict eligibility criteria, emerging variants of concern, and the changing epidemiology of the pandemic, many questions still remain unanswered regarding vaccine performance. Post-introduction vaccine effectiveness evaluations can help us to understand the vaccine's effect on reducing infection and disease when used in real-world conditions. They can also address important questions that were either not studied or were incompletely studied in the trials and that will inform evolving vaccine policy, including assessment of the duration of effectiveness; effectiveness in key subpopulations, such as the very old or immunocompromised; against severe disease and death due to COVID-19; against emerging SARS-CoV-2 variants of concern; and with different vaccination schedules, such as number of doses and varying dosing intervals. WHO convened an expert panel to develop interim best practice guidance for COVID-19 vaccine effectiveness evaluations. We present a summary of the interim guidance, including discussion of different study designs, priority outcomes to evaluate, potential biases, existing surveillance platforms that can be used, and recommendations for reporting results.

* Corresponding author at: Department of Immunization, Vaccines and Biologicals, World Health Organization, 20 Avenue Appia, Geneva 1211, Switzerland.
E-mail address: patelm@who.int (M.K. Patel).

1. Introduction

Since its emergence in December 2019, SARS-CoV-2 has caused over 170 million cases of Coronavirus disease 2019 (COVID-19), and 3.5 million deaths worldwide [1]. Phase 3 randomized-controlled trials of different SARS-CoV-2 vaccines have shown efficacies ranging from 50 to 95% against symptomatic COVID-19. Review of these results by WHO and other bodies has resulted in approval for emergency use authorization/listing (EUA/EUL) for several vaccines and, between December 2020 and February 2021, >100 countries have started vaccination programs [2]. As for every new vaccine, vaccine effectiveness (VE) evaluations, which measure the reduced risk of infection or disease among vaccinated individuals attributed to vaccination under real-world conditions, should be conducted to address questions about the post-implementation performance of these vaccines and to guide future vaccination policy and strategy.

There is a critical need for post-introduction observational studies in a variety of countries and populations for a number of reasons. First, when the vaccine is rolled out to wider population groups, the overall effectiveness of the vaccine may differ from that observed in trials, particularly in geographies or sub-populations not included or underrepresented in the trials (e.g. the very frail, persons with immunodeficiencies). Second, programmatic issues, such as sub-optimal cold-chain, extended intervals between doses and incomplete vaccine schedules, may affect VE. Third, circulating variants of concern are an emerging issue and could affect VE. Additionally, because review of data from clinical trials resulted in issuance of EUA/EUL after only several months of post-vaccination follow-up, assessing longer-term protection may only be possible in VE studies. Lastly, if future COVID-19 vaccines, booster doses, or vaccination of additional populations (e.g. children, pregnant women) are conditionally approved for use based on immunogenicity results, VE studies will be necessary to confirm the inferences from the immunological data.

Features of the COVID-19 pandemic that create special challenges in evaluating VE, include the rapidly changing epidemiology, the emergence of variants of concern, biases and confounding related to time-varying risk of infection and likelihood of receiving vaccine, and the accelerated rollout of vaccines. Given these challenges, the World Health Organization (WHO) developed interim guidance on the best practices in undertaking post-introduction evaluations of COVID-19 VE [3]. The guidance document discusses critical considerations in the design, analysis, interpretation, and reporting of COVID-19 VE evaluations. It is targeted primarily for investigators and public health practitioners planning to design and undertake COVID-19 VE evaluations and for policy makers who will interpret and apply the results of such studies. Although the guidance focuses on studies to be conducted in lower- and lower-middle income settings, the principles laid out are broadly applicable everywhere. This report summarizes the issues to consider; more detailed information, and links to developed protocols, are available in the complete guidance document, available on the WHO website.

2. Deciding to conduct a VE evaluation

The decision to conduct VE evaluations should be based both on the need for country- or region-specific VE estimates to guide vaccine policy and on the available capacity to conduct rigorous VE evaluations that are likely to yield reasonably unbiased results. VE evaluations require significant planning, technical expertise, resources, and time. It is not necessary that they be conducted in every country; well-executed VE evaluations of a specific vaccine

can provide representative results for similar settings. The following criteria are suggested to be in place to conduct high-quality VE evaluations:

- *Clear public health rationale* for conducting the VE evaluation in terms of informing country-level, regional or global policy decisions.
- *Experienced epidemiologic team* to develop protocol, execute evaluation in the field, assess biases, analyze the data, and interpret the results.
- *Dedicated staffing* including experienced field team.
- *Identified sites of enrollment.*
- *Availability of reliable diagnostic tests in the study population*, preferably real-time reverse-transcription polymerase chain reaction (rRT-PCR) testing, with ideally a sensitivity $\geq 85\%$ and specificity $\geq 98\%$. Testing should be free of charge to potential participants in VE evaluations.
- *Ability to ascertain accurately the vaccination status* of participants, usually through electronic or paper records, as well as on key potential confounders such as age, ethnicity, indigenous status, health worker status and co-morbidities.
- *Data collection, management, and analytic capacity in place.* Statistician and appropriately trained epidemiologist involvement are crucial.
- *Ability to enroll enough participants* to achieve the sample size needed.
- *Data dissemination plan in place.* Willingness to report results using standardized criteria and/or share results or data for multisite analyses.
- *Funding secured* to support a rigorous evaluation.
- *Functional ethical review committee* to review protocol expeditiously, if deemed necessary according to local research determination.

3. Outcomes of COVID-19 VE evaluations

For COVID-19 VE evaluations, several major outcomes of interest can be considered. Estimating VE against death due to COVID-19 has very high public health relevance given the disease's significant mortality [1]. However, evaluating VE against COVID-19 deaths is methodologically challenging due to the difficulty of ascertaining deaths due to COVID-19 in settings without widespread ante-mortem COVID-19 testing, the ability to enroll sufficient numbers of such deaths, and the challenge of obtaining an accurate vaccination history for a deceased person.

Understanding vaccine effectiveness against severe COVID-19 is important for guiding public health and policy-setting, as severe COVID-19 has substantial repercussions on health-care systems. While several case definitions of severe disease exist, we recommend the use of one of two widely-used definitions to screen participants for enrollment. One is the WHO COVID-19 case management definition for severe or critical disease, which is an adolescent or adult with clinical signs of pneumonia (fever, cough, dyspnea, fast breathing) plus one of the following: respiratory rate > 30 breaths/minute; severe respiratory distress; SpO₂ $< 90\%$ on room air; acute respiratory distress syndrome (ARDS); sepsis; septic shock; or death, with onset in the last 10 days [4]. The other recommended definition is based on the WHO surveillance case definition for Severe Acute Respiratory Illness (SARI), which is a person with acute respiratory infection with a history of fever or measured fever of ≥ 38 °C and cough with onset within the last 10 days and requiring hospitalization [5]. Which definition to use will depend on the site of enrollment, such as building on an already existing SARI sentinel surveillance site. In addition, we rec-

ommend that regardless of which case definition is used, all the relevant variables for both definitions be collected as the actual measured values, rather than being lumped into a range if possible (e.g., percent oxygen saturation, breaths per minute, at time of admission or enrollment), so that post-hoc comparisons of evaluations can be done using the same definitions of severity.

Evaluating VE against symptomatic COVID-19 will be the most relevant outcome when comparing to the efficacy results from a clinical trial, as this has been the primary endpoint in most trials. We recommend that one of two widely-used definitions are used to screen potential participants for symptomatic disease. The first, a modification of the WHO surveillance case definition, is a person who, within the last 10 days, has had acute onset of fever and cough or acute onset of ≥ 3 of the following signs or symptoms: fever, cough, general weakness/fatigue, headache, myalgia, sore throat, coryza, dyspnea, anorexia/nausea/vomiting, diarrhea, or altered mental status [6]. The second recommended suspected case definition to use is the influenza-like illness (ILI) definition, which is a person with an acute respiratory infection with measured fever of $\geq 38^\circ\text{C}$, cough, with onset within the last 10 days [5]. As with severe disease, all the variables for both definitions of symptomatic COVID-19 should be collected so as to apply either case definition when comparing VE evaluations. Because of the multitude of reasons for seeking COVID-19 testing, such as asymptomatic screening, and travel and quarantine related testing, we advise against including persons in VE evaluations who do not meet a clinical case definition when evaluating symptomatic and severe disease outcomes.

Some vaccines have been shown to prevent both disease and infection (e.g., measles) while others prevent disease but not infection (e.g. tetanus). There is preliminary evidence that some SARS-CoV-2 vaccines reduce infection as well as probably infectiousness of those who do become infected (i.e., transmission) [7–9]. The extent of these effects will determine how much vaccines can contribute to herd immunity and thereby to reducing transmission, protecting both vaccinated and unvaccinated people to some extent from exposure to the virus. Despite the public health importance of outcomes of infection and transmission, evaluating VE for these outcomes is more challenging than for disease outcomes. Such evaluations require active follow-up of participants, and perhaps households, to test for asymptomatic infection and perhaps viral load, requiring more resources. Optimal methodologies are being developed to assess VE against infection and transmission, and these evaluations will likely be undertaken in a limited number of settings [10,11].

For all outcomes, an important question is the duration of protection. Longer observational studies, extending beyond the three months of follow-up of most clinical trials at the time of EUA/EUL, will be needed to assess if VE wanes with time since vaccination.

4. Vaccination

There are different types of COVID-19 vaccines in various stages of use and development, with different schedules, making for a complex landscape for both delivering the vaccine and for studying VE [12]. Given the differences between vaccines, a separate VE estimate should be calculated for each vaccine, except in the case where different COVID-19 vaccines are administered to the same person, in which case a VE estimate should be calculated for each regimen used. Where sample size allows, VE analyses should attempt to investigate the VE of partial vaccination, different or delayed dosing schedules, and VE against specific variants. The estimated sample size will need to be increased to allow for these different VE estimates to be calculated with confidence.

We recommend that receipt of COVID-19 vaccination be confirmed by reviewing documentation (e.g., vaccination card, immunization registers), rather than by self-report alone. Documentation allows critical information to be collected about the vaccine product (for each dose) and the timing of vaccination, as well as reducing recall bias of self-report. Lack of COVID-19 vaccination, in contrast, might only be obtained by self-report, except in circumstances where there are complete population and vaccination registers. For the analysis, a conservative approach should be taken in considering a person as protected from vaccination as from 14 days after the date of the first dose of vaccine, and, if applicable, 7–14 days after the date of the second dose. The number of days post-vaccination to use in the primary analysis should be driven by the specifications of the vaccine being evaluated.

5. Study designs

VE studies aim to emulate the results that might be obtained from a randomized trial. However, an inherent weakness of all observational evaluations is that non-randomized vaccinated and unvaccinated groups potentially differ in key characteristics, such as risk of infection and access to testing and healthcare. One of the key challenges in observational studies is to design them so as to minimize the confounding effects of differences between vaccinated and unvaccinated persons [13]. Table 1 outlines the main study designs to evaluate VE against COVID-19 disease outcomes. Regardless of which study design is used, only those persons who are eligible for vaccine should be included.

5.1. Cohort studies

Retrospective or prospective cohort studies, in which individuals with known vaccination status are followed, allow direct calculation of the incidence of infection or disease in vaccinees versus non-vaccinees, leading to estimation of the absolute reduction in the incidence of infection or disease among vaccinated persons, as well as the VE. Cohort studies could potentially be built onto existing population-based studies, such as in demographic surveillance sites, or focus on specific populations that have been prioritized for COVID-19 vaccination, such as health workers, in which follow-up can be more efficient and complete. Additionally, with accurate electronic health records that can link vaccination status to disease outcome in individuals, a cohort study could be conducted efficiently as has been done for COVID-19 in several countries already [14–16].

5.2. Case-control studies

Case-control studies, where investigators identify individuals who were diagnosed with COVID-19 (i.e., cases), and a comparison group of individuals who were not diagnosed with COVID-19 (i.e., controls), are frequently less expensive than cohort studies as the required sample size is much smaller. However, it is challenging to select the appropriate controls to be representative of the population from which cases arise, in terms of exposure to virus and vaccination coverage, often leading to biased VE results.

5.3. Test-negative design studies

The test-negative design (TND) may be the most feasible approach in most settings. This is a common method for estimating influenza and rotavirus VE due to its logistical ease and minimization of some biases [17,18]. In a health facility (either inpatient or outpatient), patients who seek care and meet a predetermined case definition based on a predefined set of symptoms/signs (ideally

Table 1
Types of Observational Studies to Measure COVID-19 Vaccine Effectiveness [3].

Type of Observational Study	Strengths	Weaknesses	Resource requirement	Comment
Cohort Studies (prospective or retrospective)	<ul style="list-style-type: none"> • Results easily communicated to policy makers and stakeholders • Can estimate burden of COVID-19 in a population and potentially measure the impact of vaccination • Easier to interpret when done early when limited vaccine supply • Can potentially be used to study asymptomatic or mildly symptomatic infections 	<ul style="list-style-type: none"> • Vaccination status difficult to determine in retrospective cohorts without good vaccination records • Rt if outcome of interest is uncommon such as severe COVID 19 • May be expensive, especially if prospective • If prospective, possible ethical dilemma in following unvaccinated persons who are recommended for vaccination 	High	Could be undertaken in certain situations such as among healthcare workers, in institutionalized settings, Health Maintenance Organizations or sentinel hospitals with electronic medical records, or in well circumscribed outbreaks
Case-Control (CaCo) Studies	<ul style="list-style-type: none"> • Efficient as requires smaller sample size, as focus on identifying cases rather than following a large population with few cases • Less expensive than cohort studies • Most people familiar with case-control design 	<ul style="list-style-type: none"> • Need to choose controls to reflect the population from which cases arise, in terms of exposure to virus and vaccination coverage • Vaccinated persons may be more likely to seek, or have access to, health care and become cases, biasing towards reduced VE • Misclassification of vaccination status greater compared to cohort studies, especially prospective cohort studies 	Moderate	Controls should be enrolled at same time as case enrolled in changing incidence setting.
Test-Negative Design (TND) Case-Control Studies	<ul style="list-style-type: none"> • Reduces bias of differences in healthcare seeking behavior and access by vaccine status • All cases and controls seek care at same facilities, potentially decreasing differences in access to vaccines and community-level confounders • Vaccination status often obtained before results of laboratory tests available, minimizing diagnostic bias • Can use existing surveillance platforms, such as those for influenza • Logistics are simplified, less resource intensive 	<ul style="list-style-type: none"> • False negative misclassification more likely than CaCo as both cases and controls have COVID-19-like illness. • Test-negative controls more likely to be tested for exacerbation of an underlying illness (e.g., COPD), that is an indication for COVID-19 vaccination leading to increased VE. • Cases and controls need to be matched or the analysis needs to be adjusted by time • Does not remove confounding from common predictors of vaccination and exposure to infection, such as being in a priority group by age or occupation 	Moderate	Probably most efficient and least biased study design for VE studies of COVID-19 disease in most settings.
Screening Method	<ul style="list-style-type: none"> • Markedly reduced expenses since relies on available coverage data and leverages ongoing disease surveillance • Do not have to collect data among non-cases since uses vaccine coverage surveys • Estimation of expected number of cases who are vaccinated (i.e., breakthrough cases) 	<ul style="list-style-type: none"> • Coverage survey data may not be representative of population from which cases are being collected (e.g. differences in healthcare access and healthcare seeking behavior) • Vaccination status may come from administrative data rather than surveys raising concerns about validity of coverage estimate • Must have vaccine status of all reported cases • Unable to adjust for individual level covariates • Defining the "neighborhood" around cut-off value for vaccination can be challenging • Potentially small sample size • Spillover vaccination among those outside cut-off • Herd protection among unvaccinated • Age cut-offs for vaccination may change rapidly depending on vaccine availability. 	Minimal	Rapid rollout makes coverage estimate moving target; disaggregation of coverage data by target populations is difficult. Could be used to determine expected number of cases among vaccinated.
Regression Discontinuity Design	<ul style="list-style-type: none"> • Minimizes selection bias as vaccine allocation is based on programmatic criterion • Minimizes temporal and geographic trends among the groups 	<ul style="list-style-type: none"> • Defining the "neighborhood" around cut-off value for vaccination can be challenging • Potentially small sample size • Spillover vaccination among those outside cut-off • Herd protection among unvaccinated • Age cut-offs for vaccination may change rapidly depending on vaccine availability. 	Moderate	

using one of those mentioned above) are enrolled in the evaluation and tested for SARS-CoV-2. Cases are those that test positive; controls are those that test negative. The TND has several advantages. First, all cases and controls have sought care at the same facilities. Hence cases and controls will generally have come from the same communities, reducing bias due to community-level variations in vaccine access and disease risk. Secondly, cases and controls have all sought care and been tested for a similar set of symptoms, reducing confounding due to differences in health-care seeking behavior or access between cases and controls, which can be a source of bias in traditional case-control studies. Third, vaccine status is typically collected and recorded at the time of specimen collection, prior to knowing the test result, reducing the likelihood of differential exposure misclassification. The TND, however, still is subject to many of the same biases as other study designs. Since everyone in a TND undergoes laboratory testing for SARS-CoV-2, there is a particular risk of misclassification of cases and controls due to lack of perfect test performance. This is more relevant for studies focused on assessing VE against severe COVID-19, as these patients tend to become severely ill after the first week of illness when viral RNA might no longer be detectable in the upper respiratory tract.

5.4. Screening method

The screening method is a pseudo-ecologic design, in which two data points are needed to calculate VE: the proportion of reported cases occurring in vaccinated persons and the vaccination coverage in the population. As such, it is relatively easy to perform and inexpensive [19]. The screening method requires valid coverage estimates corresponding precisely to the population from which cases came. It can be difficult to adjust for some potential confounders using this design, given lack of individual-level data in the population. We recommend against the use of screening method designs for estimating COVID-19 VE in the early stages of vaccine rollout when vaccine coverage is rapidly changing; it could potentially be used in defined settings where coverage is more stable.

5.5. Regression discontinuity design

The regression discontinuity design (RDD) is a quasi-experimental design that does not randomize individuals or units but leverages programmatic assignment of vaccine allocation based on a clear cut-off value [20]. In the case of COVID-19 vaccines this would likely be an age cut-off for older adults. RDD assumes there is a similar risk of disease and distribution of confounders in “a small neighbourhood” around the cut-off (e.g. 5 years above and below age cut-off) [21]. Rates of COVID-19 would be compared between those eligible for vaccination above the age cut-off and not eligible for vaccine below the age cut-off, yielding a VE estimate. The RDD has several disadvantages that are noted in Table 1.

6. Laboratory testing

We currently recommend the use of laboratory-confirmed outcomes for COVID-19 VE evaluations, ideally with rRT-PCR on samples from the upper respiratory tract, because the presentation of COVID-19 is not sufficiently specific to distinguish it clinically from other diseases. Updated guidance from WHO for ideal samples to collect and on test-performance characteristics should be reviewed before choosing a diagnostic test, particularly in light of new variants that might affect test performance [22]. Tests with imperfect sensitivity and specificity can bias a VE estimate, with poor speci-

ficity having more impact on causing invalid VE estimate [23,24]. Ideally, any test used should have at least $\geq 85\%$ sensitivity and $\geq 98\%$ specificity to reduce the risk of misclassification bias, based on simulations run using a methodology previously described for influenza VE evaluations [23]. Antibody testing should not be used as the primary method to classify participants as cases or non-cases. This is because antibody testing has limited accuracy which would lead to misclassification of participants, antibody testing cannot determine if a person was infected prior to or after vaccination, and depending on the type of vaccine administered, one might not be able to differentiate natural from vaccine-induced antibodies.

Variants of SARS-CoV-2 have arisen with multiple mutations, leading some viruses to be deemed Variants of Concern (VOCs) due to higher transmissibility, severity, or potential to evade vaccine-induced immunity. WHO has guidance on when genomic characterization (e.g. sequencing) should be conducted as part of routine surveillance [25]. Genomic characterization within the context of VE evaluations offers an opportunity to assess whether current vaccines protect against VOCs. If possible, positive specimens from all cases in VE evaluations, regardless of vaccination status, should undergo genomic characterization. If a sufficient number of cases have their samples characterized, one could determine the VE against commonly circulating variants, with particular interest against VOC. Of note, if genomic characterization of all cases is not possible, then characterizing a representative subset of positive specimens to document the circulating variants in the population will allow the results to be interpreted in light of that context.

7. Biases

Due to lack of randomization of persons to vaccination in real-world settings, all observational studies are subject to bias, and the measured VE estimate may differ from the true VE. Biases may make a vaccine appear more or less protective than it is, and the magnitude of particular biases may change during the course of a study. Proper study designs can minimize bias. Confounding is a type of bias in which a third variable is associated with both vaccination and disease, but is not in the causal chain from vaccine to disease prevention. Some potential confounders are known and can be measured and partially controlled for in the design or analysis, such as age and sex, while others are unknown and/or unmeasurable. Results of VE evaluations should always be interpreted with the potential for residual biases in mind.

For most biases, undertaking studies when vaccination coverage in the vaccine-target group is neither too low (<10%) nor too high (>90%) is recommended, as persons who get vaccinated first, or do not get vaccinated when coverage is high, tend to have different levels of risk of exposure and/or disease, resulting in greater likelihood of biases.

Table 2 outlines potential biases of COVID-19 VE studies. Some of the key biases are discussed here. Confounding can occur when a person's vaccination status is associated with their risk of being exposed to SARS-CoV-2. If vaccinated persons are those who are at increased risk, for example health workers (HWs) treating COVID-19 patients, the risk of exposure is greater, leading to decreased estimates of VE. Conversely, some people who choose not to get vaccinated might also choose not to engage in nonpharmaceutical interventions (NPIs), putting them at higher risk of infection, thereby leading to spuriously elevated VE estimates.

Health-care seeking/access bias is when people who have better access or higher tendency to utilize health-care will both be more likely to be vaccinated and also to present for care when symptomatic. This can lead to a higher proportion of vaccinated people

Table 2
Potential biases of COVID-19 vaccine effectiveness studies [3].

Bias	Description	Designs affected*	Typical Magnitude	Direction on VE estimate	Outcomes / subgroups in which VE affected	Methods to minimize bias	Comments
Care-seeking behavior/ access to care	Those more likely to get vaccine seek care more, thus more likely to be cases	CaCo, cohort	Large	Decrease	Non-severe more than severe disease	Use TND; enroll only severe patients.	TND partially addresses, but can create collider bias [32]
Care-seeking based on vaccine status	Vaccinated persons less likely to seek care/testing due to COVID-19-like illness due to perception of protection	All	Small-moderate	Increase in CaCo and cohort; decrease in TND, if vaccine confers some protection.	Non-severe more than severe disease	Smaller magnitude in TND	Might partially offset care-seeking behavior/better access bias
Collider bias [32]	Health-seeking and SARS-CoV-2 infection both lead to testing	TND	unknown	Unknown, depends on how health-seeking and infection affect testing	Non-severe more than severe disease	Limit to severe patients; limit to older adults	
Confounding other than by factors mentioned above	Occurs when there are common causes of receipt (or lack of receipt) of vaccine and risk of SARS-CoV-2 exposure	All	Unknown	Unknown (depends on direction risk of vaccination and exposure are affected)	All	Stratification, regression adjustment, or matching for potential confounders (e.g., HW occupation)	It is important to collect high quality data on potential confounding factors, particularly adherence to NPI. Example of healthy vaccinee effect
Diagnostic bias	HWs more likely to test unvaccinated persons for COVID-19	All	Varies on setting	Increases	Non-severe more than severe disease	Test all persons or a systematic random sample meeting protocol-specified case definitions	
Misclassification of the outcome	False negatives (persons with COVID-19 disease who test negative)	TND > CaCo, cohort [23]	Small	Decrease	Severe disease more affected due to later presentation for testing	Use a highly sensitive test; limit to illness onset ≤ 10 days; exclude TND controls with COVID-19-specific symptoms (e.g. loss of taste)	Rapid tests currently have lower sensitivity than PCR; If vaccination shortens shedding time, could lead to increased estimate of VE.
Misclassification of the outcome	False positives (persons without COVID-19 disease who test positive)	TND > CaCo, cohort	Small	Decrease	All	Limit to illness onset ≤ 10 days, use highly specific test, use of clinical case definition for enrollment.	Possible chronic shedder/persistent PCR positive who is ill from another cause, but likely rare; could be more problematic when incidence is high.
Misclassification of the exposure	Vaccine effect may start before/after specified cutoff for considering individual vaccinated	all	Large but can be nearly eliminated by design	Decrease	All	Exclude from primary analysis outcomes occurring in periods of ambiguous vaccine effect, e.g. 2 weeks after first dose	Particular concern for COVID-19 when rollout is fast and large proportion of follow-up time and cases will occur soon after vaccination.
Nonspecific vaccine effect	Vaccine prevents diseases for which controls seek care	TND	Small (has not been shown)	Either; depends if vaccine increases or decreases other diseases	All	Exclude controls with diseases possibly affected by COVID-19 vaccines [33]	E.g., adenovirus-vector vaccines might prevent adenovirus illness
Prior infection	If known prior SARS-CoV-2 infection, less likely to get vaccinated	All	Small-moderate (depends on seroprevalence / past incidence of infection)	Decrease	All	Sensitivity analysis excluding those with prior SARS-CoV-2 by history or lab	Assumes prior infection confers immunity. Asymptomatic prior infection could occur in risk group targeted for early vaccine (e.g. HWs)
Spurious waning	Unvaccinated individuals become immune through natural infection faster than vaccinated [34]	All	Small soon after vaccine campaign, large with increasing time since campaign	Decreases with time since vaccination	VE of duration of protection	Do VE study soon after vaccine introduction; anchoring in time of cases and controls	Occurs with “leaky” vaccine that partially protect against infection and there is high incidence of infection [35]

Table 2 (continued)

Bias	Description	Designs affected*	Typical Magnitude	Direction on VE estimate	Outcomes / subgroups in which VE affected	Methods to minimize bias	Comments
Survivorship	Unvaccinated more likely to die of COVID-19	All	Small	Decrease	Severe disease; high-risk mortality groups	Quantify percent of COVID-19 deaths in non-study population who were vaccinated. If conducting inpatient evaluation, attempt to enroll fatal cases	Refers to deaths of person before they would have chance to be enrolled in study

* Designs include traditional case-control (CaCo), test-negative design case control (TND), and cohort studies

Table 3

Potential Reasons for Vaccine Effectiveness (VE) estimates that are different from vaccine efficacy results [3].

VE estimate valid	VE estimate not valid
<ul style="list-style-type: none"> Population being studied has different VE for epidemiologic or biological reasons Vaccine mishandling Systematic error in vaccine administration Problems with vaccine batch Waning immunity resulting in lower VE Different outcome or schedule is being evaluated from clinical trial Vaccine less effective due to mutations in SARS-CoV-2 virus Contribution of vaccine associated enhanced disease (VAED) (especially severe disease outcome) Prevalence of prior infection in population different from that of efficacy study 	<ul style="list-style-type: none"> Error in implementation (e.g. enrollment of persons not meeting case definition, poor specimen collection/handling) Biases Unmeasured or incompletely controlled confounders Chance finding; more likely with small sample size

among cases, lowering the VE estimate. TNDs partially mitigate this bias since all enrolled persons have sought care.

Several biases can occur in the first couple of weeks after vaccination. Individuals who are experiencing early COVID-19 symptoms or who were recently exposed might defer vaccination, leading to an apparent protective effect in the first few days after the first dose, when the vaccine is expected to have no effect. Persons experiencing side effects from the vaccine might be more likely to seek care and testing post-vaccination, which could also lead to an early elevated VE when using the TND. Conversely, exposure misclassification might occur soon after vaccination due to a delay between the date of infection and development of symptoms and presentation for testing, creating a delay in when positive VE becomes apparent. Additionally, infection could precede vaccination and the vaccine might not protect post-exposure, leading to an apparently decreased VE soon after vaccination. In VE studies with relatively short follow-up time, these biases in the first few weeks after vaccination might have significant impact on the overall VE estimate. To ensure validity, it may be necessary to exclude from the primary analysis outcomes occurring during the periods of approximately 14 days after the first dose and 7–14 days after the second dose, as the individual’s immunization status when they were infected may be uncertain.

Spurious waning of the VE may occur if natural infection provides strong protection against reinfection whereas a vaccine only partially protects an individual against infection, a so-called leaky vaccine [26]. The measured VE will decrease with time since vaccination, as the unvaccinated group is depleted of susceptible individuals due to natural infection occurring faster than in the vaccinated group, who get partial protection against natural infection. This bias makes evaluation of the duration of protection challenging.

Prior SARS-CoV-2 infection can create both confounding and non-confounding bias. If persons were aware of having had prior SARS-CoV-2 infection they might be less likely to get vaccinated, and also less likely to get infected if prior infection confers immunity [27].

Documenting known prior SARS-CoV-2 infection among study participants might allow for exclusion of prior infection in the analysis, or a stratified analysis as has been done in the United Kingdom [7]. While baseline serological status of participants in VE evaluations can allow for secondary analyses based on serostatus, in many resource limited settings it might not be possible to obtain baseline serology on all participants in VE evaluations. Additionally, serological testing has limitations, and not everyone with a prior infection is seropositive [22]. However, the baseline seroprevalence in the population in which the evaluation is taking place, if known, can help to quantify the expected bias on VE estimates.

8. Covariates

Covariates are variables that are collected as part of the enrollment process in VE evaluations. Besides the standard demographic and clinical data, for VE studies of COVID-19 vaccines it is important to collect a history of previous SARS-CoV-2 infection, adherence to NPIs, and if the person is in a priority group for vaccination (e.g. HW). Potential confounders of COVID-19 VE evaluations include previous SARS-CoV-2 infection, access to health-care, socio-economic status, being in a priority group for vaccination and risk reduction behaviors such as mask use and social distancing. As some covariates change over time, as does disease incidence, time of onset is often an important covariate to include in analyses. Effect modifiers (i.e. different VEs in different subgroups) could be age, chronic medical conditions, or certain medications. A detailed list of covariates to collect is provided in the guidance document [3].

9. Interpretation of results from VE evaluations

The findings of any VE evaluation should be interpreted in light of the efficacy results from the clinical trials, if available. If the VE is

unexpectedly high or low compared to efficacy estimates or other published VE results, the VE estimate may be valid in the study context. However, unexpected VE estimates could also reflect methodological flaws (Table 3). Unexpected results should be followed-up with a detailed programmatic and epidemiological evaluation. However, given the complexity of these studies, all results should lead to an examination of the implementation of the evaluation and methods for analysis to ensure that case definitions were applied consistently, that case ascertainment was appropriate, that vaccination status was appropriately determined, that known confounders were controlled for and that identifiable biases did not occur, or if biases occurred that the results were interpreted taking these into account.

10. Platforms for COVID-19 VE evaluations

A potentially efficient approach to conducting COVID-19 evaluations is to build an evaluation onto an existing platform used for another purpose. Platforms that could be leveraged include influenza SARI and ILI surveillance sites, inpatient sentinel disease surveillance sites for other diseases (e.g. acute febrile illness), already existing health worker surveillance or cohorts, cohort event monitoring for COVID-19 Adverse Events Following Immunization, and administrative databases. Outbreaks, especially in well-defined populations such as prisons, long-term care facilities, boarding schools and military barracks, also could serve as settings in which to undertake VE evaluations efficiently. Both cohort and case-control studies can be undertaken in outbreak settings. [19]

11. Reporting results

WHO encourages consistent and standardized reporting of results of COVID-19 VE evaluations that include sufficient details on study participants, data collection, and analyses to enable readers to judge the validity of the study. Lack of complete reporting of key VE study elements and heterogeneity in reporting will create limitations in being able to compare across studies conducted in different settings. Without consistent reporting, pooled analyses or meta-analyses that increase power to evaluate VE will be difficult to interpret, as observed for influenza VE evaluations [28,29]. Having a standardized format for reporting will facilitate ease of interpretation for the many audiences that will be interested in COVID-19 VE studies.

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) consensus guidelines were created to aid authors in ensuring high-quality presentation of observational studies [30]. STROBE guidelines consist of a minimum set of reporting elements for observational studies, typically compiled in a checklist that authors must complete before submitting a relevant manuscript to a journal. These include descriptions of setting, dates of enrollment and follow-up, case definitions, exposure measurement, sample sizes, patients included/excluded, and key characteristics of the study participants. The STROBE guidelines provide a starting point for COVID-19 VE reporting. However, due to the unique aspects of COVID-19 epidemiology and vaccines, additional data elements for COVID-19 specific VE studies expanding upon the STROBE checklist are recommended (Table 4). Additionally, WHO encourages sharing of COVID-19 VE evaluation databases in data repositories available to the public, to encourage transparency and facilitate pooling of results [31].

12. Conclusions

VE evaluations will play an important role in answering key programmatic and policy questions related to the multiple

COVID-19 vaccines rapidly being rolled out worldwide. However, these studies are not simple and require careful planning and sufficient technical and financial resources to ensure that the evaluations are addressing relevant questions, the design is appropriate, collected data include all important confounders and are of sufficient quality, biases are minimized, data are interpreted correctly, and data are shared in a way that facilitates comparisons and promotes transparency. To promote optimal and relevant COVID-19 VE evaluations, WHO has published best practice guidance which describes the various aspects to consider when planning and conducting a VE evaluation.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The following authors have disclosures to make. Gagandeep Kang is vice chair of CEPI, is on the board of directors of Ignite Life Science Foundation, and her institution receives funding in her capacity as independent director of the MSD Wellcome Trust Hilleman Laboratories Private Limited. Claudio F. Lanata's institution has received research funding for COVID-19 vaccine studies from CureVac AG. Maïna L'Azou Jackson was an employee of Sanofi Pasteur until January 2020. Marc Lipsitch receives financial support from Bristol Myers Squibb and Sanofi Pasteur related to general COVID-19 work for presentations. He participates on the data safety monitoring board or advisory boards of Covaxx, Pfizer, Janssen, AstraZenca, One Day Sooner (all unpaid). He is a former unpaid board member of One Day Sooner. He also receives general support from the Morris Singer Foundation and the National Cancer Institute/National Institutes of Health USA. Walter A. Orenstein is a member of the Scientific Advisory Board for Moderna. Justin R. Ortiz receives honoraria and travel support for participation on scientific advisory board for Immunization for All Ages (Pfizer); honoraria and travel support for participation on scientific advisory board for Real World Evidence (Seqirus); his institution receives research funding for COVID-19 vaccine studies from NIH and Pfizer. Peter Smith is a member of the Data Safety Monitoring Board for Imperial College's trial of SARS-CoV-2 vaccine, for Curevac's trial of SARS-CoV-2 vaccine, and for Imperial College and Oxford University's human challenge trials with SARS-CoV-2. Note that as we do not discuss any specific brand of vaccine, we do not feel any of these disclosures are a conflict of interest, but in the interest of transparency we have detailed relationships with pharmaceutical and other organizations.

Acknowledgements

We would like to acknowledge Qian An for her assistance in providing input into the statistical section of the guidance document, Tamara Pilishvili for providing into various sections of the guidance document, and Michael Jackson, Kim Mulholland and Marc-Alain Widdowson for their critical review of the guidance document.

Disclosures

The following authors have disclosures to make. Gagandeep Kang is vice chair of CEPI, is on the board of directors of Ignite Life Science Foundation, and her institution receives funding in her capacity as independent director of the MSD Wellcome Trust Hilleman Laboratories Private Limited. Claudio F. Lanata's institution has received research funding for COVID-19 vaccine studies from CureVac AG. Maïna L'Azou Jackson was an employee of Sanofi Pasteur until January 2020. Marc Lipsitch receives financial support from Bristol Myers Squibb and Sanofi Pasteur related to general

Table 4
 Strengthening the reporting of observational studies in epidemiology (STROBE) checklist [30] and recommended additional elements for reporting COVID-19 vaccine effectiveness studies* [3].

Section/Topic	STROBE Item no.	STROBE	COVID-19 VE studies
TITLE AND ABSTRACT			
Title/abstract	1	<ul style="list-style-type: none"> Indicate the study's design with a commonly used term in the title or the abstract Provide in the abstract an informative and balanced summary of what was done and what was found 	<ul style="list-style-type: none"> Specify study design (e.g., case-control, TND or cohort) Report vaccine type(s), outcome, target vaccine groups evaluated, study location, VE and 95% confidence intervals
INTRODUCTION			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	<ul style="list-style-type: none"> Mention efficacy results from pivotal clinical trial that led to EUL/EUA or licensure of vaccine being studied Describe specific vaccine products in use, timeline of introduction, targeted populations and coverage, NPI measures in place in study area Describe COVID-19 epidemiology preceding and during period of study, including baseline seroprevalence in the target population if known, disease activity, and predominant variants during the study
Objectives	3	State specific objectives, including any prespecified hypotheses	<ul style="list-style-type: none"> Was study done to provide local/subpopulation VE estimates or answer global evidence gap in VE data?
METHODS			
Study design	4	Present key elements of study design early in paper	<ul style="list-style-type: none"> TND, traditional case-control, cohort, other
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	<ul style="list-style-type: none"> Describe the enrollment setting (e.g. SARI surveillance, hospitalized patients), location or region COVID-19 incidence at time of study, vaccines in use, introduction dates, and timing of rollout in target groups, NPI measures in place, and common circulating SARS-CoV-2 variants
Participants	6	Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up. Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	<ul style="list-style-type: none"> Report time period when data were collected Report specific clinical case definition used for enrollment Report definition of severity used Describe eligible study population in terms of age and vaccine target groups (e.g., HWs, chronic medical conditions) and exclusion criteria
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	<p><i>COVID-19 vaccine variables</i></p> <ul style="list-style-type: none"> Report definition for vaccination status, including exclusions based on vaccine timing (e.g., receipt of vaccine < 14 days of illness onset) and fully vs. partially vaccinated, dose interval <p><i>COVID-19 outcomes</i></p> <ul style="list-style-type: none"> Report sensitivity and specificity of diagnostic test used; if rapid antigen test, give test name and antigen target; Indicate if COVID-19 result known prior to or after enrollment. Explain how possible vaccine reactions were handled in TND studies (e.g., exclude recent vaccinees tested for possible febrile reaction to vaccine) <p><i>Covariates</i></p> <ul style="list-style-type: none"> Report covariates assessed for confounding, and if and how adjusted for Report the specific cut points used for continuous variables that are categorized (e.g. age groups). Provide the list of conditions included as "high risk" Provide the unit if time if adjusting for calendar time. Describe how prior COVID-19 infection was defined
Data sources/measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	<p><i>COVID-19 vaccine</i></p> <ul style="list-style-type: none"> Report source of vaccination data (e.g., vaccine card, medical record, registry, provider report, patient report, or some combination of the above). List the type and brand of vaccine (lot number if available). Report recommended schedule for vaccination (number of doses and time interval between doses) <p><i>COVID-19 outcomes</i></p> <ul style="list-style-type: none"> Report procedures for collection of respiratory samples and RT-PCR testing, include type of respiratory samples collected (e.g. nasal, nasopharyngeal), type of swab used (e.g. flocced), transport media (e.g. universal transport media or report if dry swabs were used) and maximum interval from onset to swab collection; Report up to how many days before enrollment a positive COVID-19 test was acceptable; Were subjects with compatible clinical illness without lab confirmation enrolled?
Bias	9	Describe any efforts to address potential sources of bias	<ul style="list-style-type: none"> Report if prior COVID-19 infection and exposure risk to COVID-19 (e.g., mask-wearing) were assessed and how handled
Study size	10	Explain how the study size was arrived at	<ul style="list-style-type: none"> Adjust sample size calculation to expected COVID-19 incidence and estimated VE from clinical trial

(continued on next page)

Table 4 (continued)

Section/Topic	STROBE Item no.	STROBE	COVID-19 VE studies
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groups were chosen and why	<ul style="list-style-type: none"> Report the specific cut points used for continuous variables that are categorized (e.g. age groups). Provide the unit of time if adjusting for calendar time
Statistical methods	12	Describe all statistical methods, including those used to control for confounding	<ul style="list-style-type: none"> Describe the specific regression method used (e.g. logistic regression) and confidence limits methodology Report the time periods for which data were analyzed and if COVID-19 was circulating throughout Specify any matching variable (e.g. time) and whether regression model accounts for matching Specify how covariates assessed for inclusion in the model and final covariates included Describe how partially vaccinated persons were handled in the analysis (e.g., one dose) Describe how data were pooled if gathered from multiple sites and measure of heterogeneity calculated
Statistical methods	12	Describe any methods used to examine subgroups and interactions Explain how missing data were addressed Cohort study—If applicable, explain how loss to follow-up was addressed Case-control study—If applicable, explain how matching of cases and controls was addressed Describe any sensitivity analyses	<ul style="list-style-type: none"> Describe any analyses of subgroups (e.g. age groups, chronic conditions, HWs) Describe interactions assessed (e.g. prior COVID-19 infection) Describe whether a complete case analysis was used or if missing data were imputed. Name the package used for imputation (e.g. ICE in Stata). In case-control studies, if more than one control group enrolled, explain rationale. For example, excluding verbal reports of vaccination; limited to positive test within 72 h of enrollment; limited to PCR + only (if rapid antigen tests included) Indicate if and where study protocol and/or study data are publicly available
Other			
RESULTS			
Participants	13	<ol style="list-style-type: none"> Report numbers of individuals at each stage of study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completed follow-up, and analyzed Give reasons for non-participation at each stage Consider use of a flow diagram 	
Descriptive data	14	<ol style="list-style-type: none"> Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders Indicate number of participants with missing data for each variable of interest Cohort study—summarize follow-up time (e.g., average and total amount) 	<ul style="list-style-type: none"> Describe percentage of each COVID-19 vaccine used in the study population Report number of participants who received only one dose of two dose schedule, and if different vaccines given for each dose Describe seroprevalence of study population, if available
Outcome data	15	Cohort study—Report numbers of outcome events or summary measures over time Case-control study—Report numbers in each exposure category, or summary measures of exposure	<ul style="list-style-type: none"> Describe number/percent of tests which were PCR, rapid antigen test, other. Report COVID-19 genomic information among vaccine failures, if available. Particularly variants of concern.
Main results	16	<ol style="list-style-type: none"> Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence intervals). Make clear which confounders were adjusted for and why they were included Report category boundaries when continuous variables were categorized If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period 	<ul style="list-style-type: none"> Report adjusted VE and 95% CI by vaccine type Report adjusted VE and 95% CI for target groups separately, if sufficient power Report heterogeneity statistics for pooled data
Other analyses	17	Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	<ul style="list-style-type: none"> Report age-stratified VE and 95% CI estimates separately Report separate VE and 95% CI among those with one dose, two doses and at least one dose COVID-19 vaccines Report separate VE and 95% CI by SARS-CoV-2 variant if sufficient power
DISCUSSION			
Key results	18	Summarize key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	<ul style="list-style-type: none"> Specifically discuss potential biases affecting COVID-19 VE studies, including health-seeking bias, misclassification bias, diagnostic bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	<ul style="list-style-type: none"> Explain potential differences in study VE from efficacy in relevant clinical trials (e.g., different target group, different outcome, immunization system factors)
Generalizability	21	Discuss the generalizability (external validity) of the study results	<ul style="list-style-type: none"> Was baseline seroprevalence different from other settings? Predominant viral variant found in other settings?

Table 4 (continued)

Section/Topic	STROBE Item no.	STROBE	COVID-19 VE studies
OTHER INFORMATION			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

* Table modified from unpublished work by the WHO Working Group on Observational Influenza Vaccine Effectiveness Reporting Standards, 2017.

COVID-19 work for presentations. He participates on the data safety monitoring board or advisory boards of Covaxx, Pfizer, Janssen, Astra-Zenca, One Day Sooner (all unpaid). He is a former unpaid board member of One Day Sooner. He also receives general support from the Morris Singer Foundation and the National Cancer Institute/National Institutes of Health USA. Walter A. Orenstein is a member of the Scientific Advisory Board for Moderna. Justin R. Ortiz receives honoraria and travel support for participation on scientific advisory board for Immunization for All Ages (Pfizer); honoraria and travel support for participation on scientific advisory board for Real World Evidence (Seqirus); his institution receives research funding for COVID-19 vaccine studies from NIH and Pfizer. Peter Smith is a member of the Data Safety Monitoring Board for Imperial College's trial of SARS-CoV-2 vaccine, for Curevac's trial of SARS-CoV-2 vaccine, and for Imperial College and Oxford University's human challenge trials with SARS-CoV-2.

Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention (CDC) nor the World Health Organization. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding

This work was supported by German Federal Ministry of Health (BMG) COVID-19 Research and development funding to WHO and by the National Cancer Institute of the National Institutes of Health under Award Number U01CA261277.

Author contribution

MKP and DF conceived and wrote the draft of the global guidance document and this manuscript. All authors contributed to the development of the content of the guidance document, and critically revised the guidance document and this manuscript. All authors have given approval to submit this document for publication.

References

- [1] World Health Organization. "WHO Coronavirus Disease (COVID-19) Dashboard." Retrieved June 3, 2021, from <https://covid19.who.int/>.
- [2] Our World in Data. "Coronavirus (COVID-19) Vaccinations." Retrieved March 10, 2021, from <https://ourworldindata.org/covid-vaccinations>.
- [3] World Health Organization. "Evaluation of COVID-19 vaccine effectiveness: Interim Guidance, 17 March 2021." Retrieved March 28, 2021, from https://www.who.int/publications/i/item/WHO-2019-nCoV-vaccine_effectiveness-measurement-2021.1.
- [4] World Health Organization. "Clinical management of COVID-19: Interim Guidance." Retrieved January 15, 2021, from <https://www.who.int/publications/i/item/clinical-management-of-covid-19>.
- [5] World Health Organization. "WHO surveillance case definitions for ILI and SARI." Retrieved March 10, 2021, from https://www.who.int/influenza/surveillance_monitoring/ili_sari_surveillance_case_definition/en/.
- [6] World Health Organization. "WHO COVID-19: Case Definitions." Retrieved January 15, 2021, from <https://apps.who.int/iris/rest/bitstreams/1322790/retrieve>.
- [7] Hall V, Foulkes S, Saie A, Andrews N, Oguti B, Charlett A, et al. Effectiveness of BNT162b2 mRNA vaccine against infection and COVID-19 vaccine coverage in healthcare workers in England, multicentre prospective cohort study (the SIREN study). *The Lancet* 2021;preprint.
- [8] Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N Engl J Med* 2021;384:403–16.
- [9] Levin-Tiefenbru M, Yelin I, Katz R, Herzel E, Golan Z, Schreiber L, et al. "Decreased SARS-CoV-2 viral load following vaccination." Retrieved March 10, 2021, from <https://www.medrxiv.org/content/10.1101/2021.02.06.21251283v1.full.pdf>.
- [10] Lipsitch M, Kahn R. "Interpreting vaccine efficacy trial results for infection and transmission." Retrieved March 10, 2021, from <https://www.medrxiv.org/content/10.1101/2021.02.25.21252415v1>.
- [11] Follmann D, Fay M. "Vaccine Efficacy at a Point in Time." Retrieved March 10, 2021, from <https://www.medrxiv.org/content/10.1101/2021.02.04.21251133v1>.
- [12] World Health Organization. "Draft landscape and tracker of COVID-19 candidate vaccines." Retrieved March 10, 2021, from <https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines>.
- [13] Garcia-Albeniz X, Hsu J, Hernan MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol* 2017;32:495–500.
- [14] Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, et al. Effectiveness of the BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass-Vaccination Setting. *N Engl J Med* 2021.
- [15] Vasileiou E, Simpson CR, Robertson C, Shi T, Kerr S, Agrawal U, et al. Effectiveness of first dose of COVID-19 vaccines against hospital admissions in Scotland : national prospective cohort study of 5.4 million people. Preprint 2021.
- [16] Pawlowski C, Lenehan P, Puranik A, Agarwal V, Venkatakrishnan AJ, Niesen MJ, et al. FDA-authorized COVID-19 vaccines are effective per real-world evidence synthesized across a multi-state health system. [preprint medrxiv] 2021.
- [17] World Health Organization. Generic protocol for monitoring impact of rotavirus vaccination on gastroenteritis disease burden and viral strains. In: Immunizations V, and Biologicals, editor. Geneva, Switzerland 2008. <https://apps.who.int/iris/handle/10665/69913>. Accessed: January 15, 2021.
- [18] World Health Organization. Evaluation of influenza vaccine effectiveness. Geneva, Switzerland 2017. <https://apps.who.int/iris/handle/10665/255203>. Accessed: January 15, 2021.
- [19] Orenstein WA, Bernier RH, Dondero TJ, Hinman AR, Marks JS, Bart KJ, et al. Field evaluation of vaccine efficacy. *Bull World Health Organ* 1985;63:1055–68.
- [20] Lopez Bernal JA, Andrews N, Amirthalingam G. The Use of Quasi-experimental Designs for Vaccine Evaluation. *Clin Infect Dis* 2019;68:1769–76.
- [21] Basta NE, Halloran ME. Evaluating the Effectiveness of Vaccines Using a Regression Discontinuity Design. *Am J Epidemiol* 2019;188:987–90.
- [22] World Health Organization. Diagnostic testing for SARS-CoV-2: Interim Guidance. Geneva, Switzerland 2020. <https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2>. Accessed: January 15, 2021.
- [23] Jackson ML, Rothman KJ. Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness. *Vaccine* 2015;33:1313–6.
- [24] De Smedt T, Merrill E, Macina D, Perez-Vilar S, Andrews N, Bollaerts K. Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. *PLoS ONE* 2018;13:e0199180.
- [25] World Health Organization. SARS-CoV-2 genomic sequencing for public health goals: Interim guidance, 8 January 2021. Geneva, Switzerland 2021. https://www.who.int/publications/i/item/WHO-2019-nCoV-genomic_sequencing-2021.1. Accessed: January 15, 2021.
- [26] Lewnard JA, Tedijanto C, Cowling BJ, Lipsitch M. Measurement of Vaccine Direct Effects Under the Test-Negative Design. *Am J Epidemiol* 2018;187:2686–97.

- [27] Lipsitch M, Kahn R, Mina MJ. Antibody testing will enhance the power and accuracy of COVID-19-prevention trials. *Nat Med* 2020;26:818–9.
- [28] Belongia EA, Simpson MD, King JP, Sundaram ME, Kelley NS, Osterholm MT, et al. Variable influenza vaccine effectiveness by subtype: a systematic review and meta-analysis of test-negative design studies. *Lancet Infect Dis* 2016;16:942–51.
- [29] Sullivan SG, Feng S, Cowling BJ. Potential of the test-negative design for measuring influenza vaccine effectiveness: a systematic review. *Expert Rev Vaccines* 2014;13:1571–91.
- [30] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;370:1453–7.
- [31] "Harvard Dataverse." Retrieved January 15, 2021, from <https://dataverse.harvard.edu/>.
- [32] Griffith GJ, Morris TT, Tudball MJ, Herbert A, Mancano G, Pike L, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun* 2020;11:5749.
- [33] Ray GT, Lewis N, Klein NP, Daley MF, Wang SV, Kulldorff M, et al. Intraseason Waning of Influenza Vaccine Effectiveness. *Clin Infect Dis*. 2019;68:1623–30.
- [34] Lipsitch M, Goldstein E, Ray GT, Fireman B. Depletion-of-susceptibles bias in influenza vaccine waning studies: how to ensure robust results. *Epidemiol Infect* 2019;147:e306.
- [35] Lipsitch M. Challenges of Vaccine Effectiveness and Waning Studies. *Clin Infect Dis* 2019;68:1631–3.