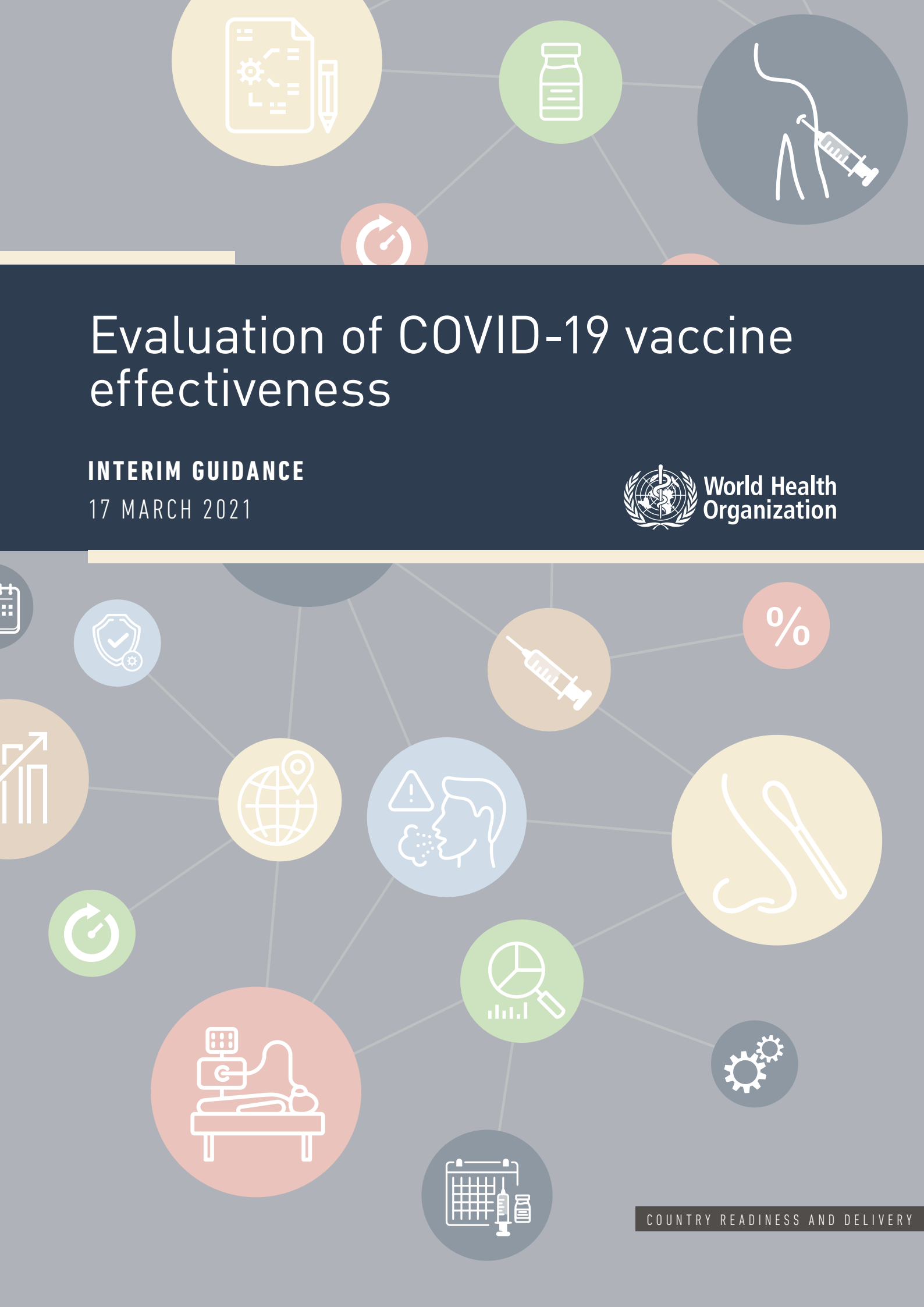
Evaluation of COVID-19 vaccine effectiveness

INTERIM GUIDANCE

17 MARCH 2021

World Health Organization



COUNTRY READINESS AND DELIVERY

Evaluation of COVID-19 vaccine effectiveness

INTERIM GUIDANCE

17 MARCH 2021

WHO continues to monitor the situation closely for any changes that may affect this interim guidance. Should any factors change, WHO will issue a further update. Otherwise, this interim guidance will expire 2 years after the date of publication.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by WHO in preference to others of a similar nature that are not mentioned.

WHO reference number: WHO/2019-nCoV/vaccine_effectiveness/measurement/2021.1

Contents

Acknowledgements

The development of this document was coordinated by Daniel Feikin and Minal K Patel of the World Health Organization (WHO) Department of Immunization, Vaccines, and Biologicals.

This guidance document was developed with technical input from:

Qian An (United States Centers for Disease Control and Prevention [US CDC])
Isabel Bergeri (Global Influenza Programme, WHO)
Joe Bresee (US CDC)
Ben Cowling (University of Hong Kong)
Natasha Crowcroft (Department of Immunization, Vaccines, and Biologicals, WHO)
Kamal Fahmy (WHO Regional Office for the Eastern Mediterranean)
Siddhivinayak Hirve (Global Influenza Programme, WHO)
Michael Jackson (Kaiser Permanente Washington Health Research Institute)
Sudhir Joshi (WHO Regional Office for South-East Asia)
Gagandeep Kang (Christian Medical College, Vellore, India)
Mark Katz (WHO Regional Office for Europe)
Claudio Lanata (Nutritional Research Institute, Peru)
Maïna L'Azou Jackson (Coalition for Epidemic Preparedness Innovation)
Marc Lipsitch (Harvard University, United States of America)
Jason Mwenda (WHO Regional Office for Africa)
Francisco Nogareda (Consultant to Pan American Health Organization)
Walt Orenstein (Emory University, United States of America)
Justin Ortiz (University of Maryland, United States of America)
Richard Pebody (WHO Regional Office for Europe)
Tamara Pilishvili (US CDC)
Alba Maria Ropero (Pan American Health Organization)
Stephanie Schrag (US CDC)
Peter Smith (London School of Hygiene and Tropical Medicine, United Kingdom of Great Britain and Northern Ireland)
Padmini Srikantiah (Bill & Melinda Gates Foundation)
Lorenzo Subissi (Emerging Diseases and Zoonoses, WHO)
Marta Valenciano (Epiconcept)
David Vaughn (Bill & Melinda Gates Foundation)
Jennifer Verani (US CDC)
Annelies Wilder-Smith (Department of Immunization, Vaccines, and Biologicals, WHO)

Kim Mulholland (Royal Children's Hospital, Melbourne, Australia, and London School of Hygiene and Tropical Medicine, United Kingdom) and Marc-Alain Widdowson (Institute of Tropical Medicine, Antwerp, Belgium) gave a critical review of the document.

Abbreviations

AEFI	adverse events following immunization
aOR	adjusted odds ratio
ARDS	acute respiratory distress syndrome
aRR	adjusted relative risk
ARU	attack rate among the unvaccinated
ARV	attack rate among the vaccinated
CaCo	case-control study
CEM	cohort event monitoring
CEPI	Coalition for Epidemic Preparedness and Innovations
CI	confidence interval
CLIA	chemiluminescence immunoassays
COPD	chronic obstructive pulmonary disease
COVID-19	coronavirus disease 2019
DBP	diastolic blood pressure
ECMO	extracorporeal membrane oxygenation
ELISA	enzyme-linked immunosorbent assays
ERC	ethical review committee
EUA	Emergency Use Authorization
EUL	Emergency Use Listing
Hib	*Haemophilus influenzae* type b
HMO	health maintenance organization
ICU	intensive care unit
ILI	influenza-like illness
IVIR-AC	Immunization and Vaccine-related Implementation Research Advisory Committee (WHO)
LFI	lateral flow immunoassays
L/MICs	low- and middle-income countries
LRT	lower respiratory tract
NPI	non-pharmaceutical interventions
RDD	regression discontinuity design
rRT-PCR	real-time reverse-transcription polymerase chain reaction
RSV	respiratory syncytial virus
SAGE	Strategic Advisory Group of Experts on Immunization (WHO)
SARI	severe acute respiratory infection
SBP	systolic blood pressure
SES	socioeconomic status
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TND	test-negative design case-control
URT	upper respiratory tract
US CDC	United States Centers for Disease Control and Prevention
VAED	vaccine-associated enhanced disease
VE	vaccine effectiveness
WHO	World Health Organization

Key messages

This document provides interim best practice guidance on how to assess COVID-19 vaccine effectiveness (VE) using observational study designs. It discusses critical considerations in the design, analysis and interpretation of COVID-19 VE evaluations, as biased results may be produced even in settings where data completeness and quality are high. This guidance is targeted mostly for evaluations undertaken in low- and middle-income countries (L/MICs), but most of the concepts apply to VE evaluations in high-income settings as well.

Key messages in this document:

- Due to their methodological complexity and susceptibility to biases, COVID-19 VE evaluations do not need to be conducted by all countries introducing COVID-19 vaccines. A checklist of criteria to have in place when considering such evaluations is provided.

- Objectives of VE evaluations are to evaluate real-world performance of vaccines, to address gaps in evidence from clinical trials (including effectiveness in subgroups, effectiveness against variants of concern and duration of protection), to provide input into impact models, and to provide post-authorization confirmation of effectiveness of conditionally approved products.

- The most feasible outcomes to evaluate in VE evaluations in most settings are symptomatic disease and severe disease. VE studies of death, infection and transmission, while of great public health importance, generally require targeted special studies with more resources.

- We recommend the use of laboratory-confirmed outcomes in VE evaluations. At this time, we recommend the use of real-time reverse-transcription polymerase chain reaction (rRT-PCR) for laboratory testing of participants. Specimens should be taken within 10 days of disease onset.

- We recommend against the use of self-reported COVID-19 vaccination as the sole source indicating whether a person is vaccinated, due to recall bias and lack of product details. Documented vaccination should be used for the primary analysis; self-reported vaccination could be included in a secondary analysis.

- Although not without potential biases, we recommend the test-negative design as the most efficient and logistically feasible method to assess VE in L/MICs, with the advantage of some degree of comparability between cases and controls since they all sought care for a similar illness at the same facilities. Other methods that could be considered are cohort studies, case-control studies, and the screening method (in certain settings with reliable information on coverage at different times during the study period).

- Due to lack of randomization of vaccination in real-world settings, all observational study designs are subject to bias because vaccinated persons often differ from unvaccinated persons in their disease risk, independent of vaccination. Important biases include the following: confounding by health care seeking or exposure risk, misclassification of outcomes due to diagnostic errors, prior SARS-CoV-2 infection, and spurious inferences of waning. Collection of key covariates to control for confounding bias in the analysis is important.

- For the primary analysis of VE studies, a conservative approach is recommended in considering a person as potentially protected from vaccination only from 14 days after the date of first dose of vaccination (the time required to achieve protection for the majority of vaccine recipients for most vaccines), and 7–14 days after second doses of vaccine (if applicable).

- The primary analysis should compare persons receiving the recommended number of doses of the same vaccine with unvaccinated individuals. Secondary analyses include partially vaccinated persons, persons receiving doses of two different vaccines, targeted subgroups, viral variants, and history of prior SARS-CoV-2 infection or disease if available. Even though partial vaccination and the use of different vaccines to complete a course are not currently recommended by WHO, these might happen in the real world and findings could inform future policy.

- VE estimates that vary from efficacy in clinical trials could be valid or not valid; thorough investigation into reasons for the difference should be undertaken.

- Existing platforms that can be used for VE evaluations include surveillance systems for severe acute respiratory infections (SARI), influenza-like illness (ILI), other syndromic disease surveillance in sentinel hospitals, health worker surveillance, administrative databases and well-defined outbreaks.

- We recommend standardized reporting of the results of studies based on the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidance, as well as suggested additional COVID-19 specific elements described below.

1. Introduction

1.1 Purpose of this guide and target audience

Since its emergence in December 2019, SARS-CoV-2, the virus that causes coronavirus disease 2019 (COVID-19), has taken a tremendous toll globally; by 28 February 2021, there have been over 110 million cases and 2.5 million deaths worldwide from COVID-19 *(1)*. Although most COVID-19 deaths occur among older adults and persons with chronic comorbid medical conditions, deaths have occurred in persons of all ages. Moreover, the pandemic has caused widespread morbidity and necessitated control measures that have devastated economies worldwide. In response to the pandemic, the global efforts to develop multiple vaccines to protect against COVID-19 disease have been unrivalled in the history of public health. By the end of 2020, three COVID-19 vaccines have received Emergency Use Approval/Listing (EUA/EUL) by maturity level 4 regulatory authorities, based on reaching predefined criteria for safety and efficacy, and at least several dozen more are in clinical trials *(2)*.

From December 2020, vaccines started to be rolled out according to various allocation plans, which differ by country. Generally, these are based on criteria of risk of serious disease and death, ethical principles of fairness and equity, and considerations for restarting stalled economies *(3, 4)*. As vaccine production capacity scales up and new products are authorized, allocation criteria will broaden until supply enables widespread use of vaccines.

During the initial implementation phases, as for every new vaccine, post-introduction evaluations will be important to address many of the remaining questions about the performance of these vaccines. When a vaccine is used outside trial populations the effects of the vaccine may differ in specific geographies or subpopulations. Vaccine effectiveness (VE) might be different against various disease outcomes, against infection and infectiousness, and against newly emerging virus variant strains. Additionally, important programmatic issues will need to be addressed, such as the effectiveness of incomplete dose schedules, variation in dose intervals, and the interchangeability of different vaccine products. Suboptimal cold chain capacity, and off-schedule and incomplete delivery of doses could lead to different vaccine performance. Vaccines might not be as effective against new variants. Finally, assessing the duration of vaccine protection requires longer term studies. This document offers best practice guidance in how to undertake post-introduction evaluations of the effectiveness of COVID-19 vaccines.

Most clinical trial results are likely to be from high- and middle-income country populations and assessing effectiveness in representative low- and lower middle-income countries (L/MICs) will be particularly important. This document emphasizes approaches deemed most feasible in L/MICs.

This guidance for undertaking VE evaluations for COVID-19 vaccines follows, in many aspects, previous World Health Organization (WHO) guidance on how to evaluate VE in observational studies, including for vaccines against rotavirus, influenza and *Haemophilus influenzae* type b (Hib) and pneumococcus *(5–7)*. Because of its similarity in clinical presentation and epidemiology, this guidance document draws heavily on the influenza VE guidance. Nonetheless, several distinct features of COVID-19 epidemiology and vaccines create unique challenges and approaches to evaluation.

This guidance is written primarily for investigators and public health practitioners who will design and undertake observational COVID-19 VE evaluations and for policy-makers who will interpret and apply the

results of these evaluations. The document discusses critical considerations in the design, analysis and interpretation of COVID-19 VE evaluations, as biased results may be produced even in settings where data completeness and quality are high. These recommendations also aim to ensure a level of comparability and completeness of reporting on studies that will enable comparability between studies.

Importantly, we do not recommend that COVID-19 VE evaluations be conducted by all countries introducing COVID-19 vaccines. VE evaluations will likely be conducted by a number of countries worldwide, for a variety of different vaccines, and the results can be expected to be applicable to other countries in the same region with similar populations, COVID-19 epidemiology and immunization systems. The decision to carry out VE evaluations should be based on the need for country- or region-specific VE estimates to guide vaccine policy and on capacity to conduct rigorous VE evaluations that will minimize biases and optimize the likelihood of accurate results.

1.2 Epidemiology of COVID-19

COVID-19 was first identified in Wuhan, China in December 2019. WHO declared a public health emergency of international concern on 30 January 2020 and a pandemic on 11 March 2020. Knowledge about transmission of the SARS-CoV-2 virus, the causative virus, is continuously evolving as new evidence accumulates. According to available evidence, SARS-CoV-2 mainly spreads between people when an infected person is in close contact with another person *(8,9)*. The estimated incubation period (from infection to onset of first symptoms) is between 2 and 14 days with a median of 5 days. Importantly, infection may also be asymptomatic but still be transmissible *(10–12)*.

A wide range of symptoms for COVID-19 has been reported, the most common being acute onset of fever, chills, cough and shortness of breath. Loss of smell or taste are symptoms that seem to be more common than with other viral respiratory infections. The majority of SARS-CoV-2 infections are either asymptomatic or result in mild disease. Some persons will develop post-acute COVID-19 syndrome (also known as "long COVID-19"). Data from several countries suggest that 14%–19% of ill persons are hospitalized, and 3%–5% will develop severe disease that requires intensive care unit (ICU) admission for complications *(13–15)*. Radiologic findings, consisting mostly of ground-glass patterns in the lung parenchyma, have been found even among mildly symptomatic persons. Older age, particularly > 60 years, is the strongest risk factor for severe disease and death *(16–18)*. Underlying noncommunicable diseases, such as diabetes, hypertension, cardiac disease, chronic lung disease and cancer, are also risk factors for severe disease and death *(16, 19–22)*. Some racial groups, morbidly obese persons and pregnant women also have elevated risk for severe disease *(23–26)*.

Clinical manifestations of COVID-19 are generally milder in children than in adults. However, a rare acute presentation with a hyperinflammatory syndrome leading to multiorgan failure and shock temporally associated with COVID-19 has been described as multisystem inflammatory syndrome in children *(27)*.

1.3 COVID-19 vaccine landscape, regulatory status and policy considerations

With unprecedented speed, by the end of 2020, over 200 vaccine candidates on various platforms were in development, of which 14 are in late clinical stage development, and three have received EUA/EUL by maturity level 4 regulatory authorities and have started to be rolled out in multiple countries. Additional vaccines have received national regulatory approval and are in use in a few countries, some in advance of the results of efficacy trials *(2)*. Many countries have set up bilateral agreements with multiple manufacturers to procure COVID-19 vaccines *(28)*. However, most L/MICs have relied on the COVAX Global

Vaccine Facility, which is co-led by WHO, the Coalition for Epidemic Preparedness and Innovations (CEPI) and Gavi, the Vaccine Alliance *(29)*. Within the COVAX Facility, resources secure the R&D of its vaccine portfolio and pooled demand for equitable access to COVID-19 vaccines for all participating countries. Through the COVAX Facility, countries will receive vaccine allocations proportional to their population size which will be deployed according to national allocation frameworks *(30)*.

All COVID-19 vaccines for in-country use will need to be authorized by national regulatory authorities. Most of the early approvals for use are likely to be based on interim results of efficacy trials, leading to EUA/EUL or conditional approval. EUA/EUL is not licensure, and formal licensure must still be obtained for these vaccines. After EUA/EUL by regulatory authorities, COVID-19 vaccines can be submitted to WHO's vaccine prequalification programme. WHO also has an EUL mechanism for COVID-19 vaccines (31). Approved vaccines for purchase through the COVAX Facility require WHO prequalification, or the EUA/EUL of a maturity level 4 regulatory authority (e.g. the United States Food & Drug Administration, European Medicines Agency).

The Strategic Advisory Group of Experts on Immunization (SAGE) advises WHO on vaccine policy. SAGE reviews evidence from the COVID-19 vaccine clinical trials and provides product-specific recommendations, particularly with respect to use of vaccines in L/MICs. Of note, not all COVID-19 vaccines will be submitted for WHO prequalification, be recommended for use by SAGE, nor be procured through the COVAX Facility. Some vaccines might be used in countries through bilateral arrangements with manufacturers. It is important that these vaccines also be evaluated by well executed post-introduction VE evaluations in settings in which they are rolled out.

1.4 Suggested criteria to undertake VE evaluations of COVID-19 vaccines

VE evaluations require significant planning, technical expertise, resources and time. The following are criteria suggested to be in place to conduct a high-quality VE evaluation:

- ☐ **Clear public health rationale** for conducting the VE evaluation in terms of informing policy decisions in country, region or globally. Participation of ministries of health in VE evaluation is encouraged to facilitate use of the data to guide policy.

- ☐ **Experienced epidemiologic team** to develop the protocol, execute evaluation in the field, assess biases, analyse the data and interpret the results. Consultation of technical partners with experience in VE studies is recommended.

- ☐ **Dedicated staffing,** including experienced field team, to enrol participants by rigorously applying screening case definitions and ensure any necessary testing, to complete questionnaires and conduct follow-up, as needed for some study designs; personnel at each enrolment site will likely be needed. Other key staff include supervisors, data specialists and administrative support.

- ☐ **Identified sites of enrolment:** existing surveillance platforms are an advantage but are not a necessary prerequisite for conducting studies. However, if new surveillance platforms are set up, this will take time to establish. Electronic medical record databases might also be used if available.

- ☐ **Availability of reliable diagnostic tests in the study population,** preferably real-time reverse-transcription polymerase chain reaction (rRT-PCR) testing, with ideally a sensitivity ≥ 85% and specificity ≥ 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.

- ☐ **Data collection, management and analytic capacity in place:** statistician and appropriately trained epidemiologist involvement are crucial. Understanding potential sources of bias and having the ability to accurately capture data on potential confounding variables are needed.

- ☐ **Ability to enrol enough participants** to achieve the required sample size needed. The time to complete VE evaluations depends on many variables (e.g. outcomes, design, incidence), but usually requires several months at a minimum.

- ☐ **Data dissemination plan in place:** willingness to report results using standardized criteria and/or share results or data for multisite analyses.

- ☐ **Funding secured to conduct rigorous evaluation:** costs will vary depending on country costs, existence of platforms that can be leveraged, study design, and sample size. Funding may be needed for additional staffing, lab collection and testing supplies, transport, or data management equipment.

- ☐ **Functional ethical review committee** to review protocol expeditiously, if deemed necessary according to local research determination.

2. The role of VE and impact studies for COVID-19 vaccines

2.1 The role of VE studies in evaluating COVID-19 vaccines

While the decision to introduce proven effective COVID-19 vaccines might not be in doubt for most countries due to the disease's high public health and economic burden worldwide, understanding these vaccines' effectiveness in real-world settings will still be essential post-introduction. The Phase III clinical trials will not answer all of the questions of the performance of these vaccines. Most of the trials are using symptomatic SARS-CoV-2 as their primary endpoint. Due to the current high incidence, many trials will achieve (or have already achieved) the primary objective after a few months and manufacturers will then apply to regulators for EUA/EUL. When the EUA/EUL is approved, these vaccines will become available quickly since manufacturing has already begun, as has been observed with the first few vaccines. Many placebo recipients in these clinical trials will likely receive active vaccine before the end of the trial, thereby limiting the ability to achieve the statistical power to evaluate many of the secondary outcomes. Some of these secondary outcomes will be critical for policy-makers, such as the efficacy against severe COVID-19 and death, the duration of protection, and risk factors for vaccine failure. Moreover, to date, the trials did not target enrolment of some groups that might be vaccinated, such as pregnant women.

In the future, COVID-19 vaccines might be conditionally authorized for use based on immunogenicity data, if this is thought to be a likely surrogate measure of efficacy, and such vaccines will require post-authorization evidence of VE against disease. Key questions to be answered about COVID-19 vaccines will likely only be answered through observational studies of VE after they are introduced (Box 1).

BOX 1. OBJECTIVES OF POST-INTRODUCTION COVID-19 VE EVALUATIONS

1. **Evaluate real-world performance of vaccines rather than in the carefully controlled conditions of a trial:**
   - **for example, cold chain (especially ultra-cold storage), timing and completeness of dosing schedule, general population including persons excluded from participation in the trials, and different circulating virus variants.**

2. **Address gaps in evidence of vaccine efficacy from clinical trials:**
   - **outcomes of interest (e.g. severe disease, death, symptomatic or asymptomatic infection, transmission)**
   - **subpopulations at risk (e.g. the very old, persons living with HIV)**
   - **duration of protection from vaccines (e.g. will there be a need for revaccination?)**
   - **whether new variants or antigenic drift of virus will affect VE**
   - **effectiveness of vaccines co-administered with routine vaccines.**

3. **Provide input into models that estimate impact of vaccines on health and economic indicators**

4. **Provide post-authorization confirmation of effectiveness of conditionally approved products for regulatory bodies**

2.2 COVID–19 vaccine effectiveness versus impact

BOX 2. MEASURES OF VACCINE PERFORMANCE

Vaccine efficacy: reduced risk of infection or disease among vaccinated individuals resulting from vaccination in carefully controlled circumstances; estimated from randomized clinical trials.

Vaccine effectiveness: reduced risk of infection or disease among vaccinated individuals attributed to vaccination in real-world conditions; estimated from observational (non-randomized) studies.

Vaccine impact: reduction in incidence of infection or disease in a population where some members are vaccinated. Vaccine impact depends upon vaccine coverage and results from direct effects of vaccination in the vaccinated, as well as any indirect effects in the vaccinated and unvaccinated due to herd protection *(32, 33)***. Impact can also pertain to other measures besides disease, such as health systems' functioning and capacity and economic indicators.**

Evaluations of how vaccines work in a population typically consider three parameters – efficacy, effectiveness and impact (Box 2). Efficacy is often estimated in pre-licensure clinical trials. This guidance document will focus mostly on real-world VE. However, a brief description of COVID-19 vaccine impact is warranted, as the reduction in the overall incidence of disease in the population due to a vaccination programme gives a key measure of the effect on public health. The impact of vaccination programmes on disease burden is typically evaluated using surveillance systems that compare the incidence of disease before and after vaccine implementation. For certain vaccines, these assessments can include both the incidence of laboratory-confirmed outcomes, such as invasive *S. pneumoniae* or rotavirus disease, and of non-specific outcomes, such as hospitalizations and/or deaths from pneumonia or diarrhoea, with the difference in incidence attributed to the vaccination programme. These pre-post impact studies, often applying interrupted time-series methods to surveillance data, can show the degree to which introduction of the vaccine has reduced disease incidence *(34)*. Pre-post impact studies, however, are most suitable for endemic diseases with fairly consistent year-to-year epidemiology, such as Hib, pneumococcus and rotavirus. When disease rates and severity vary substantially year to year, such as with influenza, pre-post impact studies are usually difficult to interpret. In the case of COVID-19, there will be little more than 1 year of baseline incidence data in the pre-vaccine period in most places. Moreover, the spread of SARS-CoV-2 has been variable, impacted by different non-pharmaceutical interventions (NPI) over time and geographical areas, with peaks and troughs of disease that vary both temporally and spatially, precluding pre-post impact evaluations in most settings. Additionally, care seeking and COVID-19 diagnostic patterns might change after vaccines are rolled out, as has been seen with other vaccine introductions. For these reasons, formal pre-post impact studies, such as interrupted time series, are likely to be challenging for COVID-19 vaccines. If pre-post impact studies are undertaken, measures to assess temporal trends would be important. For example, investigators could make a comparison of contemporaneous changes in COVID-19 rates between groups or areas targeted for vaccines and groups not targeted, or assess changes in testing rates or percentage positive among tested persons after vaccine introduction *(35)*. Nonetheless, in many settings, such studies will be very difficult to interpret, unless there are clearly evident reductions in incidence coincident with vaccine introduction.

Vaccine impact could also be assessed by a phased introduction of vaccine by geographic area (e.g. stepped wedge design), which might allow for contemporaneous assessment of the reduction in incidence due to widespread vaccination. This might be an attractive approach in the setting of limited supply of vaccine. Limitations to this approach include variable COVID-19 epidemiology in different geographies

and rapid vaccine rollout leaving an insufficient amount of time to evaluate rates between vaccinated and unvaccinated areas. Moreover, it could be politically unpalatable, unethical, impractical or counter-productive to delay vaccination in certain areas.

Apart from rigorously designed special studies, we do not recommend undertaking vaccine impact studies in most L/MIC settings, especially in short time intervals after vaccine introduction. Interpretation of pre-post impact data should be done with caution.

3. Outcomes of interest for VE evaluations

For COVID-19 VE evaluations, several major outcomes of interest can be considered.

3.1 COVID-19 deaths

The burden of COVID-19 mortality has been devastating, with 1.7 million deaths occurring in the first year of the pandemic and the number still increasing rapidly *(1)*. Preventing COVID-19 deaths has been one of the principles that has guided recommendations for early COVID-19 vaccine allocation, though none of the controlled trials have been large enough to evaluate efficacy against this endpoint. Therefore, measuring VE against COVID-19 deaths would have very high public health relevance.

VE evaluations of COVID-19 deaths, however, are challenging methodologically. First, many people who have died from COVID-19 have not undergone diagnostic testing, making it difficult to distinguish COVID-19 and non-COVID-19 deaths. This might be particularly the case in L/MIC settings, when many cases may not seek hospital care before dying. Attempts to identify COVID-19 deaths retrospectively using verbal autopsy are likely to be inaccurate due to the lack of specificity in defining COVID-19 based on symptoms and signs. Second, it will be difficult to accumulate enough confirmed COVID-19 deaths in a study site, since the incidence of COVID-19 mortality remains relatively low, especially in some L/MIC settings where there are fewer older adults *(36)*. Third, obtaining accurate vaccination status for deceased people may be difficult in most L/MICs. **Because of these concerns, in most settings, we do not recommend VE evaluations where the primary outcome is only COVID-19 mortality.**

Importantly, however, deaths are a subset of those with severe disease (see Section 3.2 Severe COVID-19 disease( )) . In some settings, some assessment of vaccine impact on COVID-19 deaths might be possible. Some settings might have large enough administrative databases (e.g. health maintenance organizations [HMOs], electronic health records in large hospitals) with standardized cause of death determinations and vaccine status that a retrospective analysis of COVID-19 VE against death could be possible. If highly effective vaccines become available in a short period of time with high coverage among high-risk groups, a reduction in excess mortality ascribed to COVID-19 might reflect a large vaccine impact (though will not necessarily estimate VE) *(37)*.

3.2 Severe COVID-19 disease

Severe disease is an outcome of public health and policy setting relevance as it is a precursor to death, can result in long-term morbidity, and has substantial repercussions on health care systems. Confirming severe COVID-19 disease is likely achievable in many settings. Hospitalization can serve as a reasonable minimal criterion for severity in most settings, which will allow for logistical efficiency in capturing severe cases. Yet, health care utilization and thresholds for hospitalization can vary by geographic location and individual hospital, leading to variability of severity among hospitalized cases. Therefore, hospitalization alone should not be used as the case definition for severe disease, as it is not comparable across settings.

Several definitions of severe disease, including those used in clinical trials, can be found in Annex 3: Possible case definitions; inclusion and exclusion ( ). The definitions used in the Phase III clinical trials

and the Solidarity treatment trial rely on variables that might only be available in advanced hospital settings, such as level of ventilatory support. For L/MICs, we recommend the use of one of two more widely applicable definitions to screen for cases.

1. WHO COVID-19 case management definition for severe or critical disease:

   Adolescent or adult with clinical signs of pneumonia (fever, cough, dyspnoea, fast breathing) plus one of the following: respiratory rate > 30 breaths/min; severe respiratory distress; or SpO2 < 90% on room air; or acute respiratory distress syndrome (ARDS), sepsis, septic shock, or death *(38)*. (Death not included in case management definitions but can be included in VE evaluation definition.)

2. WHO surveillance case definition for Severe Acute Respiratory Illness Infection (SARI):

   A hospitalized 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 *(39)*.

Which definitions to use will depend on the site of enrolment; the case management definition does not require admission to hospital. In addition, we recommend that regardless of which case definition is used, all the relevant variables for both definitions be collected, as measured values, rather than in categories, where possible (e.g. percentage oxygen saturation, breaths per minute, at time of admission or enrolment), so that post-hoc comparisons of evaluations can be done using the same definitions of severity.

3.3 Symptomatic COVID-19 disease

Symptomatic SARS-CoV-2, i.e. COVID-19, is the primary outcome of most vaccine clinical trials, usually including severe and non-severe disease.

Many people with mild symptoms of COVID-19 disease will not present for medical care; to capture all such cases for VE evaluations would require frequent active follow-up. However, some percentage of symptomatic cases will present for health care; this could be in outpatient settings or hospitals. In most places, it will be more feasible logistically to focus enrolment among medically attended cases. Nonetheless, because the decision to seek health care as well as the decision to test someone for COVID-19 at a health facility, can be influenced by multiple factors, including vaccination status, bias can be introduced (see Section 7. Bias in VE studies of COVID-19 vaccines()). To minimize this bias, we suggest that screened participants should meet a minimal set of clinical criteria suggestive of an acute respiratory tract infection. We suggest that L/MICs choose either the WHO surveillance case definition *(40)* or the influenza-like illness (ILI) case definition *(39)*, especially if using a pre-existing ILI platform.

Modified WHO surveillance case definition:
A person who has had the following symptoms within the last 10 days:
- ☐ Acute onset of fever AND cough; OR
- ☐ Acute onset of ANY THREE OR MORE of the following signs or symptoms: fever, cough, general weakness/fatigue, headache, myalgia, sore throat, coryza, dyspnoea, anorexia/nausea/vomiting, diarrhoea, altered mental status.

ILI case definition:
- ☐ A person with an acute respiratory infection with measured fever of ≥ 38°C, cough, with onset within the last 10 days.

As with severe disease, all the variables for both definitions should be collected so as to apply either case definition when comparing VE evaluations.

3.4 COVID–19 Infection and transmission

Some vaccines have been shown to prevent disease and infection (e.g. measles), while others prevent disease, but not infection (e.g. tetanus). A related question to if vaccines prevent COVID-19 infections is whether they decrease infectiousness among people who become infected. The extent to which the vaccine can reduce infection and infectiousness 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. Also, if vaccines can reduce infection, it would indicate that vaccination of children, who do not frequently experience severe disease, but who may participate in transmission of the virus, may be important from a public health standpoint.

Despite the public health significance, evaluating VE against infection and transmission will be more difficult than for disease outcomes. Evaluations of VE against infection are logistically challenging, expensive, and will likely require long-term active testing of cohorts of vaccine recipients and non-vaccine recipients, regardless of symptoms, either for detection of virus or antibodies to indicate exposure. While this can be done as part of a clinical trial in an unbiased way, VE evaluations of infection in observational studies can be subject to the same biases as VE evaluations for disease outcomes *(41)*. Moreover, following a cohort of unvaccinated individuals over time might pose ethical issues, especially after vaccine supply has improved.

Short of formal studies, early indications of the effects on transmission will be provided by evaluation of the change in epidemiology of disease in groups that are not vaccinated (particularly those living or working in close proximity to those vaccinated). Reduction of transmission might be observed through phased introductions of vaccine that show reduction in disease among both vaccinated and unvaccinated persons in vaccinated areas compared with unvaccinated areas. Evaluation of infection rates among households with and without vaccinated household members can be a targeted and efficient approach to address the VE against transmission. A generic standardized WHO protocol exists for household transmission studies and can be adapted to estimate the VE against transmission *(42)*. However, these can be challenging to undertake especially in light of widespread community transmission and/or multiple vaccinated persons per household *(43)*.

Despite the need for VE evaluations of infection and transmission, it is unlikely that these will be feasible in many settings, at least early in the vaccine rollout when vaccine coverage is likely to be low in the general population. Moreover, due to the need for active follow-up and testing among cohorts or households, VE evaluations of infection and transmission will be several-fold more costly and logistically complicated than those of disease. **VE evaluations of infection and transmission have critical public health significance, but we suggest that these should be undertaken as targeted special studies in a limited number of places that have adequate resources and infrastructure to undertake studies of this type.**

3.5 Laboratory-confirmed versus syndromic outcomes

In general, laboratory-confirmed outcomes have substantial advantages over syndromic outcomes for COVID-19 VE evaluations. Laboratory-confirmed outcomes are much more specific for COVID-19 disease than are syndromic outcomes based on clinical signs and symptoms. Even during periods of high incidence of COVID-19 disease, other pathogens can cause similar respiratory illness syndromes, including other

coronaviruses, influenza, respiratory syncytial virus (RSV), para-influenza, and bacterial causes. Because syndromic outcomes will be less specific for COVID-19, VE estimates based on syndromic outcomes will always be lower compared with those using laboratory-confirmed outcomes. Additionally, given the potential that vaccines are less effective against some SARS-CoV-2 variants, laboratory-confirmed outcomes (with genomic characterization) may allow for calculation of VE estimates against specific variants if the sample size is sufficient (see Section 8. Laboratory ()). In addition, all Phase III clinical trials have used laboratory-confirmed outcomes; therefore, if VE evaluations are being done to compare efficacy from clinical trials with real-world VE, then it will be important to make that comparison using laboratory-confirmed outcomes. For these reasons, **we recommend the use of laboratory-confirmed outcomes for COVID-19 VE evaluations, ideally with rRT-PCR**.

3.6 Duration of protection

One of the most important questions about COVID-19 vaccines is for how long they will prevent disease and/or infection. Because clinical trials might only follow-up participants for a few months for efficacy outcomes, assessing longer duration of protection of these vaccines will likely require post-licensure VE evaluations.

Clinical trials will assess primary vaccine failure, i.e. what percentage of individuals are not protected against disease soon after vaccination. Secondary vaccine failure can be considered as a waning of protection over time *(44)*. For some vaccines, waning can be detected by serological surveys that show decreases in protective antibody concentrations over time. These surveys are more useful from a policy perspective when there is an established correlate of protection; however, at this time, no correlate of protection is known for COVID-19. Moreover, as with some other vaccines (e.g. hepatitis vaccines) some degree of protection might be due to cellular immunity, which might vary depending on vaccine type and dose. The role of immunological markers in assessing waning of protection due to COVID-19 vaccines needs to be further studied.

Several biases can affect VE evaluations of duration of protection, particularly spurious waning of VE due to differential rates of infection and depletion of susceptible persons between vaccinated and unvaccinated cohorts (see Section 7. Bias in VE studies of COVID-19 vaccines ()). Bias of the VE over time can also occur if either health seeking or diagnostic testing changes over time, or with age, and is related to vaccination status.

There are several ways to try to distinguish true waning of VE from biased estimates of waning.

1. Look for increases in the absolute number or absolute rate of breakthrough cases among only vaccinated individuals with increasing time since vaccination for a specific vaccine product. This might give an initial signal that protection is waning, though care should be taken as the number of breakthrough infections would also increase as transmission increases. It is helpful to collect information about severity of disease as an indicator of whether vaccine modifies disease severity, and if so whether severity of disease in breakthrough cases increases over time since vaccination. It is also critical to collect information on the circulating strains as this could impact VE over time and confound the trend in breakthrough cases. Outbreak investigations that examine the relationship between time since vaccination among cases can facilitate this approach, as has been done for other vaccine-preventable diseases *(45)*.

2. Related to 1 above, follow a cohort of only vaccinated individuals and compare rates of disease in strata defined by duration since date of vaccination. Temporal and geographic trends in incidence of disease

in the general population can confound this and should be controlled for in the analysis *(45)*. If there are individuals with varying dates of vaccination all during a period of low incidence, this comparison can be especially informative as it makes the depletion-of-susceptibles bias less plausible *(46)*.

3. Optimize the study design of VE evaluations using standardized case definitions and rigorous collection of potential confounders, particularly health care seeking behaviour and clinician testing bias, to allow adjustment in the analytic phase. In addition, data on potential indicators of high exposure or susceptibility should be collected and included in the analysis *(47)*. This will also enhance the ability to combine data from different studies and perform meta-analyses of duration of protection.

4. Biases of waning VE can affect different control groups differentially. Using different methods, such as both a test-negative design (TND) and a case-control design in which controls are frequency matched by date of enrolment to allow better approximation of true waning *(48)*.

5. Groups of vaccinated and unvaccinated can be "anchored" in time. For example, matching vaccinated and unvaccinated persons in age and time of follow-up since vaccination of the matched vaccinated person, and then analyse VE by strata of follow-up time. Anchoring in time, however, can still be affected to some degree by depletion-of-susceptibles bias.

6. Use advanced statistical analytic methods that are able to better adjust for time-dependent variables that could confound the VE estimate over time *(49–51)*.

4. Assessing COVID-19 vaccination history

As part of an observational COVID-19 VE evaluation, each subject's COVID-19 vaccination history must be ascertained. The ultimate goal is to assess whether the subject has received a vaccine prior to disease onset, which vaccine, how many doses, and when. This means that the vaccination status, the vaccine product (for each dose), and the timing of vaccination need to be determined, by recording the number, product, and dates of vaccinations. Timing of vaccination is important to make sure subjects are counted as vaccinated only *after* they have received vaccine (with a time period to allow the vaccine to take effect); various time periods after vaccination can be considered as immunologically protected in the analysis.

Vaccination history can be assessed in several ways:

- One approach is to rely on administrative records created by vaccination programmes, such as patients' vaccination cards. As part of the COVID-19 vaccine rollout, countries are recommended to issue standardized vaccination cards to all persons who are vaccinated.

- Another approach is to assess COVID-19 vaccination status through either electronic or paper vaccination registries. Where available, these registries can greatly facilitate the determination of the vaccination status of participants. However, in some settings it may be difficult to match a participant to a record in the registry and data entry errors can lead to incorrect information in the registries. Registry recorded vaccine receipt has high specificity but variable sensitivity depending on quality controls in place.

If a person reports not being vaccinated, then such data can be accepted from self-report, but ideally, a search of vaccination registries should be attempted to confirm the absence of vaccination. If there is no documentation of vaccination but self-report of having received vaccine, then a secondary analysis can be undertaken in which the verbal report of vaccination is accepted. Note that persons providing self-reported vaccination history will be unlikely to know which vaccine product(s) was provided, challenging the calculation of product-specific VE in countries where multiple vaccine products are in use.

Due to limitations of recall bias, the lack of product information and dates of vaccination, **we recommend against the use of self-reported COVID-19 vaccination as the sole source indicating that a person is vaccinated.**

At a minimum, COVID-19 vaccination programmes should ensure a systematic way to record vaccine history so that the information is easily accessible *(52)*. If record keeping related to COVID-19 vaccinations can be designed with subsequent observational studies in mind, the likelihood of successfully executing those observational studies will be much greater.

At a minimum, the following data should be obtained for each participant:

- **Has the person been vaccinated?**
- **How many doses of vaccine has the person received?**
- **Is this based on patient/parental recall or documentation (e.g. vaccination card, immunization registry, medical records)?**
- **What is the date of SARS-CoV-2 vaccine dose 1?**
- **What is the product name of SARS-CoV-2 vaccine dose 1?**
- **What is the date of SARS-CoV-2 vaccine dose 2?**
- **What is the product name of SARS-CoV-2 vaccine dose 2?**

Ideally, one would pilot methods for collecting vaccine history before beginning any VE evaluation. This pilot testing should assess the accuracy and completeness of any proposed method for measuring vaccination history, and protocols should be adapted accordingly. In some cases, incorporating multiple approaches for assessing vaccination status may be the best way to ensure that vaccine information is complete and accurate.

5. Measuring covariates

Covariates are variables that are collected as part of the enrolment process into VE evaluations, which are separate from those that define the exposure (i.e. vaccination) and the outcome (i.e. COVID-19 disease/infection). Table 1 lists important covariates to consider collecting. The table distinguishes covariates that are deemed as core variables to be collected in all VE evaluations, from those considered optional; some covariates are core for some study designs only. Standardized variables and related questionnaires for some of those covariates are proposed by WHO as part of the Unity Studies standardization initiative *(53)*.

There are several reasons to collect information on covariates. First, they facilitate description of the population studied, and allow for assessment of the similarities of characteristics between groups being compared (i.e. vaccinated vs unvaccinated, those with and without COVID-19). An important reason for a descriptive analysis is to characterize the study population's risk of COVID-19. Second, some covariates will be confounders or effect modifiers of the VE estimates. Confounders are those variables that are related to both COVID-19 and vaccination status, and not in the causal pathway between vaccination and disease prevention, which can potentially distort the VE estimate. Potential confounders of COVID-19 VE evaluations include previous SARS-CoV-2 infection, access to health care, socioeconomic status (SES), and risk reduction behaviours such as mask use and social distancing. Effect modifiers are variables that define subgroups in which VE could be truly different. Effect modifiers could be age, chronic medical conditions, or certain medications. At this early stage of VE evaluations, it is not clear whether some covariates will be potential confounders, effect modifiers, or both.

Table 1. Important covariates to consider collecting

Evaluation population/ setting	Category of variable	Variable	Core/ optional	Comments
All	**Time**	• Date of illness onset, date of specimen collection[a]	Core	• Vaccination coverage and incidence change over time • Allows time-matched analysis or adjustment by time • May be critical to make comparisons stratified by date of onset if both vaccination coverage and incidence show time trends within the study period, to avoid confounding
		• Date of illness onset, date of specimen collection[a]	Core	• To assess if VE may vary with time since vaccination (waning, antigenic drift of circulating variants) • To determine if a person has sufficient time since vaccination for an immune response to develop
	Susceptibility to infection	• Previous SARS-CoV-2 infection[a]	Core	• Vaccine effect may be different among previously infected • Previous infection might be known or unknown; antibody testing of participants might be available in some settings • Previously infected may be more/less likely to be vaccinated, more/less likely to be exposed to the virus, and less likely to be infected again due to acquired immunity • Document date of and how a person's previous infection was diagnosed (e.g. laboratory-confirmed by rRT-PCR or rapid test, epidemiologically linked, or clinical)
	Variables defining priority groups for vaccination	• Health worker • Older adult	Core	• Vaccination priority groups associated with vaccination and with risk of outcome, especially severe outcomes • VE may vary according to priority group for vaccination

Table 1. Important covariates to consider collecting continued

Evaluation population/ setting	Category of variable	Variable	Core/ optional	Comments
All continued	**Sociodemographic**	• Age[a]	Core	
		• Sex[a]	Core	
		• Sociodemographic group (e.g. race[a], ethnicity, religion)	Optional (but recommended)	• Sociodemographic group, such as race may be strong predictors of access to/acceptance of vaccination as well as incidence, thus important confounders
		• Proxy for socioeconomic status (SES)	Core	• Important in settings where infection risk and probability of vaccine receipt vary by SES stratum
		• Occupation[a]	Optional	• May predict both priority group for vaccine and risk of infection/ exposure and thus be a confounder
	Chronic medical conditions	• Pre-existing chronic conditions[a]	Core	• Some chronic conditions define priority groups for vaccination due to risk of severe disease (e.g. chronic obstructive pulmonary disease [COPD], heart disease, chronic renal disease, diabetes) and may also be predictive of outcomes (via greater precautions by persons with comorbidities, or via greater risk of severe outcomes if exposed) *(54)* • Other conditions might be important to document due to potential for lower VE (e.g. HIV, other immunosuppressive disorders)
	Chronic medical conditions and medications before vaccination or disease	• Hospitalizations for chronic conditions in previous months	Core for TND	• For TND only: test-negative controls may have been hospitalized for non-infectious exacerbation of their chronic condition and therefore more likely to be vaccinated *(55)*
		• Functional impairment/frailty	Optional	• Can use questions from existing frailty index (e.g. Barthel index) • Frailty is an indication for early vaccination and is a risk factor for severe disease, though some (e.g. those at the end of their life) might be less likely to get vaccinated
		• Medications before vaccination or disease	Optional	• Important to collect in first evaluations to understand the effect on VE • Important for VE evaluations of severe disease outcome
	Access to health care	• Number of previous primary care visits (or other proxy variables)[a]	Optional	• Use of health care services associated with likelihood to be infected and to be vaccinated • TND minimizes this selection bias • To minimize recall bias, ask about past 3 months
	Other vaccinations	• Influenza, pneumococcus *(56)*	Optional	• Important to collect in first evaluations; little data on co-vaccination from clinical trials available; history of other vaccines can be marker for access to health care
	Smoking status[a]		Optional	• Collect in first evaluations to understand how smoking affects VE • Consider asking if "never, former, current smoker"
	Other respiratory viruses at time of disease	• Influenza, RSV, adenovirus	Core for TND	• For TND evaluations in particular, if COVID-19 vaccine has an effect on other respiratory viruses can create bias • Potential for use as bias-indicator/negative control outcome (see Section 7. Bias in VE studies of COVID-19 vaccines ( ))

Table 1. Important covariates to consider collecting continued

Evaluation population/ setting	Category of variable	Variable	Core/ optional	Comments
All continued	**Exposures to SARS-CoV-2 in the community**[a]	• Contact with confirmed or suspect cases within last 14 days • No. of household members • Use of public transport • Recent attendance at social events/ gatherings • Recent travel	Optional	• To characterize viral exposure risk among participants • To adjust for differential exposure to the virus between vaccinated and unvaccinated • Variables collected will depend on the setting • Adapt questions to local context, for example, number of times attended social gatherings of > 10 people
Setting specific	**Adherence to non-pharmaceutical interventions (NPIs)**	• Use of masks, adherence to social distancing, etc.	Core	• To adjust for differential risk-taking behaviour between people who choose to get vaccinated and those who do not • Persons who are adhering to NPIs might be more likely to be vaccinated • Adapt questions to local context, for example, quantification and/ or type of mask wearing when indoors outside of home during the past 2 weeks (e.g. always, mostly, sometimes, rarely, never)
	Exposures in the health care facility or long-term care facility[a]	• Health worker category • Aerosol-generating procedures? • Number of COVID-19 patients contacted and average time per patient • Personal protective equipment use	Core for health worker evaluation	• To adjust for differential exposure to the virus between vaccinated and unvaccinated

Note: [a] Standardized ways of naming and asking about these covariates in field questionnaires are available in the WHO Unity generic protocols *(53)*.

6. Study designs

Observational VE studies aim to emulate a randomized trial, in which vaccinated and unvaccinated individuals are comparable in their likelihood of being exposed to the virus and experiencing the outcome, apart from the key difference of whether they have received the vaccine. Observational studies cannot guarantee this comparability because vaccination is not randomly assigned, but they can attempt to approximate it using a variety of designs *(57)*. The common weakness of all observational evaluations is that vaccinated and unvaccinated groups potentially differ in key characteristics, such as risk of infection and access to testing, and measures should be taken to minimize this difference in all designs. Each design has strengths and weaknesses, and some designs are better for evaluating certain populations and in particular settings. Some study designs are less amenable to most L/MIC settings due to constraints of existing infrastructure and resource requirements. Ideally, planning for the evaluation should begin prior to implementation of a vaccination programme, as performing VE evaluations soon after vaccine rollout often gives the least biased estimate of VE. In many target age groups, vaccination coverage can quickly become very high making most VE designs challenging. In addition, consulting investigators with prior field experience in designing and implementing these evaluations will greatly facilitate the implementation of evaluations. Table 2 outlines the main study designs to evaluate VE against COVID-19 disease outcomes.

Table 2. Types of observational studies to measure COVID-19 vaccine effectiveness

Type of observational study	Strengths	Weaknesses	Resource requirement	Comment
Cohort studies (prospective or retrospective)	• 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	• Vaccination status difficult to determine in retrospective cohorts without good vaccination records • Requires large sample size, especially 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 health care workers, institutionalized settings, HMOs or sentinel hospitals with electronic medical records, or in well circumscribed outbreaks
Case-control (CaCo) studies	• 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 CaCo design	• Need to choose controls who 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 with cohort studies, especially prospective cohort studies	Moderate	Controls should be enrolled at same time as cases enrolled in changing incidence settings

Table 2. Types of observational studies to measure COVID-19 vaccine effectiveness continued

Type of observational study	Strengths	Weaknesses	Resource requirement	Comment
Test-negative design (TND) case-control studies	• Reduces bias of differences in health care seeking behaviour 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	• False-negative misclassification more likely than in 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	• 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)	• Coverage survey data may not be representative of population from which cases are being collected (e.g. differences in health care access and health care seeking behaviour) • 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	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	• Minimizes selection bias as vaccine allocation is based on programmatic criterion • Minimizes temporal and geographic trends among the groups	• Defining the "neighbourhood" 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	

6.1 Cohort studies

Cohort studies follow a population, with known (though possibly changing over time) vaccination status on each day, over a period of time. Disease incidence is calculated among vaccinated and unvaccinated persons. The cohort method can be used either prospectively before disease has occurred, or retrospectively using historical data in which disease has already happened. In both prospective and retrospective cohorts, vaccination status at the time of disease occurrence is determined. The cohort design allows direct calculation of the rate of disease in vaccine recipients versus non-vaccine recipients, leading to estimation of risk reduction of disease among vaccinated persons. A person can contribute both unvaccinated and vaccinated person-time, if vaccinated during the course of follow-up. Despite these advantages, cohort designs require a large sample size and are resource intensive as follow-up needs to be conducted serially for a duration of time. Moreover, vaccinated persons often differ in key characteristics from unvaccinated persons, leading to biases, and these patterns may change over time

during the course of a prioritized vaccine rollout; though this bias can be addressed, at least in part, by matching and assessed by calculation of VE for negative control outcomes (e.g. time period or for a disease when no effectiveness expected) *(56)*.

Due to these constraints, cohort studies should only be taken in a limited number of settings with capacity for rigorous follow-up to determine vaccine and disease status and adequately adjust for confounding (e.g. accurate and complete covariates). Cohorts can potentially be built onto existing population-based studies, such as demographic surveillance sites, or focus on specific vaccine-targeted populations in which follow-up can be more efficient and complete (e.g. health workers). Lastly, if vaccine is already recommended for a given group, it could be unethical to observe unvaccinated persons in that group over time without actively facilitating vaccination. However, in practice, some persons might choose not to be vaccinated, or there might be a gradual programmatic rollout of vaccines, and this group can ethically be included as non-vaccine recipients.

6.2 Case–control studies

In the case-control design, 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). Various approaches to selecting controls from the underlying population have been proposed; concurrent enrolment of cases and controls at the same time is encouraged to minimize time-variant exposure differences (58). COVID-19 vaccination history is then determined for all cases and controls, and the odds of vaccination in each group is calculated. Control enrolment should be concurrent to cases by time (e.g. onset of symptoms or testing) in a high-incidence setting, as with COVID-19 in most settings currently.

Advantages of the case-control design include efficiency in terms of cost and time to conduct the evaluation and the opportunity to address other parameters of interest (e.g. VE of less than full series of doses, duration of protection). Case-control studies of VE are inherently problematic in finding an unbiased control group that is comparable to cases in most characteristics apart from disease status, and which is not affected in some way by vaccination. Specifically, one needs to choose controls who have similar viral exposure risk and vaccination coverage to the population from which cases come. Both hospital-based and community controls have unique biases and challenges in enrolling an unbiased comparison group (59, 60). Lastly, COVID-19 should not be in the causal pathway of the reason for hospitalization when using hospitalized controls (e.g. COPD exacerbation caused by COVID-19).

6.3 Test–negative case–control design (TND)

This is a widely used method for estimating influenza and rotavirus VE due to its logistical ease and minimization of some biases *(6, 7)*. It is most often viewed as a variant of the case-control design but can also be considered as a cohort design in which all individuals who do not meet the clinical case definition that prompts testing (usually the vast majority) are lost to follow-up. In a health facility (either inpatients or outpatients), patients who seek care for a defined set of symptoms/signs are enrolled in the evaluation and tested for SARS-CoV-2. Cases are those that test positive; controls are those that test negative. It is possible that an existing surveillance platform, such as SARI or ILI surveillance, can be leveraged for COVID-19 VE evaluations (see Section 10. Platforms to do COVID-19 VE evaluations ( )). 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. This reduces confounding due to differences in health care seeking behaviour or access between cases and controls, which is often a source of bias in traditional case-control studies, particularly in outpatient settings where care seeking can be more variable. 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.

Given the current burden of COVID-19 and the low levels of other circulating viruses, it might be challenging to enrol sufficient numbers of test-negative controls in this design *(61, 62)*. Thus, investigations should consider a second set of controls (persons admitted for non-COVID-like illness who also test negative) to ensure sufficient controls are enrolled.

One particular concern of the TND is that misclassification of cases and controls can occur due to lack of perfect test performance. This is particularly relevant for VE evaluations of severe COVID-19 disease, 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 (URT). However, even with subpar sensitivity, in the context of near-perfect specificity such as provided by PCR testing, false negatives will have a negligible impact on VE estimates derived by TND (63). Additionally, if there is concern that persons present later in the course of their illness, one could consider adding exclusionary criteria for those persons who test negative, but who are highly likely to have COVID-19 *(40)*. For example, if a person tests negative, but is found to have recent onset of anosmia or ageusia or chest imaging showing findings suggestive of COVID-19, this person could be excluded from enrolment retrospectively (or excluded in a sensitivity analysis). The TND can also be subject to the same biases as other study designs, and the TND introduces some new biases, such as collider bias *(64)*. Several ways to mitigate the biases of TND are outlined in the section on bias (see Section 7. Bias in VE studies of COVID-19 vaccines ( )). **Although not without biases, we recommend conducting the TND as an efficient and accurate method in L/MICs to assess VE against severe and symptomatic COVID-19.**

6.4 The screening method (case–population method)

The screening method is a pseudo-ecologic design, which uses individual-level data on vaccination history from cases, and vaccination coverage in the source population from which the cases came. It is an attractive method in settings where disease surveillance data are available, but where few other resources are available as it does not require ascertainment of vaccination status of non-cases. Only two data points are needed to calculate VE: the proportion of reported cases occurring in vaccinated persons, which can be calculated from surveillance data; and the vaccination coverage in the population, which may be estimated from vaccine coverage surveys or available from a national registry or administrative databases. As such, it is relatively easy to perform and inexpensive *(65)*.

The screening method requires valid coverage estimates corresponding precisely to the population from which cases came. For COVID-19 vaccines, this is unlikely to be available in the first year as denominators for many of the target populations are unclear and coverage would need to be available by each targeted group. Moreover, coverage will likely be rapidly changing in the accelerated vaccine rollout phase. It can be difficult to adjust for some potential confounders using this design, given lack of individual-level data in the population. Studies using the screening method should only be undertaken in settings where vaccine coverage is stable and can be measured with high accuracy. **We therefore 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.**

Note that the screening method was originally developed to serve as a "screening" tool to understand if the proportion of cases who have received vaccine are within the expected range, or if there is a need for more rigorous investigation. For example, if a vaccine has a true VE of 70%, at 75% coverage in the population, then it would be expected that approximately half of the COVID-19 cases would be vaccinated; at 90% VE and 90% coverage, one would also expect about half of the cases to be vaccinated *(65)*. Therefore, the screening method serves as a useful tool to determine if the number of vaccine breakthorough cases is within the expected range.

6.5 Regression discontinuity design (RDD)

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 *(66)*. In the case of COVID-19 vaccines this would likely be an age cut-off for older adults or for adolescents in whom the vaccine is not authorized, versus slightly older individuals in whom it is. 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) *(67)*. 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, allowing for calculation of a rate ratio and a VE estimate. The advantages of the RDD are that it minimizes selection bias in self-determined vaccine allocation and minimizes temporal and geographic trends among the comparison groups. However, RDD has several disadvantages including how to define the number of years around cut-off before groups begin to differ, small sample size in the "neighbourhood" around cut-off, spillover vaccination if the vaccination cut-off is not adhered to, herd protection among unvaccinated persons (e.g. spouses of different ages), and rapid age de-escalation of vaccine rollout. Moreover, in many L/MICs, older persons might not know their exact age so applying strict age cut-offs for vaccination will be challenging. While RDD is a powerful design, it is likely suitable in few settings where its limitations can be addressed.

7. Bias in VE studies of COVID-19 vaccines

Due to lack of randomization of persons to vaccination in real-world settings, observational studies are subject to bias, which leads to a systematic deviation of the VE found in studies from the true VE. Biases may cause deviations in either direction making the vaccine look more or less protective than it is, and the magnitude of particular biases may change during the course of a study. 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 potentially controlled for in the analysis, while others are unknown and/or unmeasurable. Selection bias can occur when criteria for inclusion in the analysis (for example, not having had a documented prior infection with COVID-19) induce non-comparability between those who have and have not been vaccinated *(68, 69)*. Selection bias is harder to control for in the analysis and is better addressed in the design and execution of the study. All observational studies – and indeed many randomized studies – are susceptible to bias. It is important to recognize that while some designs and analytic methods can minimize biases, none can fully eliminate them from observational studies. Nonetheless, this does not mean that such studies should not be undertaken; rather efforts should be made to minimize biases, and the results should be interpreted with the potential for residual biases in mind.

For most biases, undertaking studies when vaccine 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. In areas with rapid uptake of vaccines in target populations, to avoid being left with a biased comparison group, the evaluation will likely need to be done over a several month period. Of note, highly effective vaccines might make the impact of bias less consequential, as VE estimates post-introduction will likely still be high enough to not lead to changes in existing vaccine policy.

Table 3 outlines potential biases of COVID-19 VE studies, their magnitude and direction, as well as ways to minimize them.

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

Bias	Description	Designs affected[a]	Typical magnitude	Direction on VE estimate	Outcomes/ subgroups in which VE affected	Methods to minimize bias	Comments
Care seeking behaviour/ 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; enrol only severe patients	TND partially addresses, but can create collider bias *(64)*
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 behaviour/ better access bias

Table 3. Potential biases of COVID-19 vaccine effectiveness studies continued

Bias	Description	Designs affected[a]	Typical magnitude	Direction on VE estimate	Outcomes/ subgroups in which VE affected	Methods to minimize bias	Comments
Collider bias *(64)*	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. health worker occupation)	It is important to collect high-quality data on potential confounding factors, particularly adherence to NPI. Example of healthy vaccine recipient effect
Diagnostic bias	Health workers 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 *(63)*	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 enrolment	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 cut-off 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

Table 3. Potential biases of COVID-19 vaccine effectiveness studies continued

Bias	Description	Designs affected[a]	Typical magnitude	Direction on VE estimate	Outcomes/ subgroups in which VE affected	Methods to minimize bias	Comments
Non-specific 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 *(70)*	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 infection by history or lab	Assumes prior infection confers immunity; asymptomatic prior infection could occur in risk group targeted for early vaccine (e.g. health workers)
Spurious waning	Unvaccinated individuals become immune through natural infection faster than vaccinated (46)	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 *(71)*
Survivorship	Unvaccinated more likely to die of COVID-19	All	Small	Decrease	Severe disease; high-risk mortality groups	Quantify percentage of COVID-19 deaths in non-study population who were vaccinated; if conducting inpatient evaluation, attempt to enrol fatal cases	Refers to deaths of person before they would have chance to be enrolled in study

Note: [a] Designs include traditional case-control (CaCo), test-negative design case control (TND), and cohort studies.

Some of the more relevant biases for COVID-19 VE evaluations are further described here.

- **Confounding:** This can occur when a person's vaccination status is associated with their risk of being exposed to SARS-CoV-2. This bias can lead to either spuriously high or low VE estimates. If vaccinated persons are those who are at increased risk, for example, health workers 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 NPIs, putting them at higher risk of infection, thereby leading to spuriously elevated VE estimates. Importantly, this form of bias can occur even in TND. Confounding may be a particularly important concern for VE studies in highly effective vaccination campaigns, which rapidly reach a large number of people in targeted groups, leaving those unvaccinated (or vaccinated late) either very different from those within their targeted groups who do get vaccinated (for example, members of an older age group with little social connection or access to information) or leaving mainly members of non-targeted (e.g. young healthy) groups unvaccinated. Careful statistical control for such confounding is likely to be very important to minimize bias in such circumstances.

- **Health care seeking/access:** People who have better access or higher tendency to utilize health care will both be more likely to get vaccinated and present for care when symptomatic, including with COVID-19. In traditional case-control studies, this would lead to over-representation of vaccinated individuals as cases, which would decrease the VE estimate. TNDs partially mitigate this bias since all enrolled persons have sought care. However, TNDs can lead to collider bias, whereby health seeking and SARS-CoV-2 infection both lead to testing, which is usually thought to be of lower magnitude than health care seeking bias *(64)*.

- **Diagnostic bias:** Clinicians might be less likely to order COVID-19 testing in vaccinated patients, reasoning that vaccinated patients are protected against COVID-19. TND partially addresses this bias since all participants are tested. The decision to test potential study subjects should not be based on clinicians' decisions but on prespecified protocol-defined criteria. These criteria should then be applied to all (or to a random sample of) eligible patients regardless of clinical testing decisions.

- **Misclassification of the outcome:** Outcome misclassification occurs due to imperfect laboratory test performance in diagnosing COVID-19 infection *(72)*. Erroneous test results can be both false negative and false positive. Because both rRT-PCR and rapid antigen tests tend to have higher specificity than sensitivity, false-negative test results are more common. However, false-positive tests can result in greater bias in estimating the VE *(63, 73)*. Misclassification can bias the VE more in TND than in cohort or traditional case-control studies because the control group in TND will be over-represented with false-negative cases compared with the source population. Despite these concerns, misclassification bias is likely small when using tests with high analytic sensitivity and specificity (see Section 8. Laboratory ( )). This bias can be reduced by excluding test-negative controls with COVID-19 specific symptoms (e.g. loss of taste or smell), and limiting participants to those with more recent symptom onset (e.g. ≤ 10 days) to minimize the chance of enrolling true COVID-19 cases who present late when viral RNA is no longer detectable. Another approach could be to have a second control group that is not enrolled as part of a TND, but is derived from an alternate source, such as patients hospitalized with non-respiratory illness (see Section 6.2 Case-control studies ( )). Lastly, a "bias-indicator" or "negative-control outcome" study that measures the VE of COVID-19 vaccines against a sham outcome not caused by SARS-CoV-2 can help quantify misclassification bias *(70, 72, 74, 75)*.

- **Exposure misclassification:** Given that disease outcomes are logged on the date of a positive test rather than on the date at which an individual becomes infected, the effect of the vaccine may be observed only after the incubation period plus an additional delay from symptom onset to the individual's test data, a delay that may vary by setting and health care availability. To ensure validity, it may be necessary to exclude from the primary analysis outcomes occurring during the period 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. This comes at a cost, particularly when studying the VE of the first dose of a two-dose series, because it may lead to excluding more than half the time between the first and second dose. However, failing to do so may include outcomes against which the vaccine could not have protected the individual, thereby reducing the estimate of VE from its true value (see Section 9.3.5 Time since vaccination ( )).

- **Spurious waning:** This occurs if a vaccine only partially protects against infection, a so-called leaky vaccine *(76)*. The measured VE will decrease with time since vaccination, as the unvaccinated group is depleted of non-immunes due to natural infection faster than in the vaccinated group, who get partial protection against natural infection. This is more likely to occur in settings with a high force of infection from SARS-CoV-2 during the follow-up period. This can affect all types of observational (and randomized) studies, especially when some infections are unobserved because they are subclinical and/or when exposure or infection risk is heterogeneous *(46)*. VE estimates are more likely to be valid soon after vaccination begins. For assessment of duration of protection, several approaches are discussed elsewhere (see Section 3.6 Duration of protection ( )).

- **Prior infection:** Prior SARS-CoV-2 infection can create both confounding and non-confounding bias, and VE could be different among previously infected persons compared with uninfected persons. 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 disease if prior infection confers immunity as some data suggests *(77)*. This would create confounding, and a downward bias that could lead to lower or in extreme cases even negative estimates of VE. Documenting known prior SARS-CoV-2 infection among study participants might allow for a stratified analysis or exclusion for prior infection in the analysis, as would be preferable in settings where policy discourages or prohibits those with known prior infection from getting vaccinated. Unknown prior infection, such as asymptomatic infection, in contrast, is unlikely to influence a person's likelihood of getting vaccinated. Such infection probably still confers some protection against subsequent disease. Therefore, asymptomatic infection is associated with disease, but not vaccination, and is not a true confounder. Yet, unknown prior infection can still lead to biased estimates of VE in cohort studies, in which VE is biased towards the null, compared with the value it would have if evaluated in only unexposed individuals. The bias of unknown prior infection is likely minimal in case-control studies, as it should occur in similar proportions of cases and controls. While baseline serological status of participants in VE evaluations can allow for secondary analyses, in most L/MIC settings it will not be possible to obtain baseline serology on all participants in VE evaluations, and is not a requirement to undertake such evaluations. 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. Laboratory considerations

Laboratory-confirmed outcomes are more specific for SARS-CoV-2 infection than are outcomes based on clinical signs and symptoms, which can overlap with other acute respiratory infection etiologies (see Section 3.5 Laboratory-confirmed versus syndromic outcomes ( )). Accurate laboratory results from highly sensitive and specific tests decrease the probability of misclassification. Thus, **we recommend the use of laboratory-confirmed outcomes in VE evaluations.**

However, the use of laboratory-confirmed outcomes in COVID-19 VE evaluations requires appropriate laboratory capacity, including specimen collection, handling and storage, as well as molecular assay technology and reagents. All of these activities require training of clinical staff to ensure standardized specimen collection and handling, as well as training of laboratory staff to ensure proper processing and testing of specimens.

Full details of specimen collection and laboratory testing are available from WHO *(78)*.

BOX 3. BACKGROUND INFORMATION ON SARS-COV-2 RNA DETECTION

Once an individual has been exposed to the virus, the virus may be detectable in the upper respiratory tract (URT) 1–3 days before the onset of symptoms. The concentration of SARS-CoV-2 in the URT is highest around the time of symptom onset, after which it gradually declines *(78)*. The presence of viral RNA in the lower respiratory tract (LRT), and for a subset of individuals in the faeces, increases during the second week of illness *(79)*. In some patients the viral RNA may only be detectable for several days, while in other patients it can be detected for several weeks, possibly months, although prolonged viral RNA detection beyond a week rarely indicates the presence of live virus or ongoing infectiousness.

8.1 Testing for confirmatory diagnosis

8.1.1 Laboratory methods

The laboratory test used should be sensitive and highly specific. Nucleic acid amplification testing, such as rRT-PCR, is the standard test for laboratory confirmation of SARS-CoV-2 infection during acute illness, with a high analytic sensitivity and specificity *(80)*. At this time, **we recommend the use of rRT-PCR for laboratory testing of participants in VE evaluations.**

Guidance from WHO on test performance characteristics are constantly being updated and should be reviewed to ensure the currently circulating SARS-CoV-2 viruses can still be detected by the testing modality chosen for the VE evaluation as new variants could lower sensitivity or specificity of the test *(78)*. Persons with indeterminate results should have the specimen retested, and if needed, the patient should have a second specimen collected. Although other assays are available, including rapid antigen tests, these are not ideal for VE evaluations due to issues of lower sensitivity and/or specificity at this time. Tests with imperfect sensitivity and specificity bias a VE estimate, with specificity having more impact on the VE estimate *(63, 73)*. Ideally, any test used should have at least ≥ 85% sensitivity and ≥ 98% specificity to minimize the risk of misclassification bias. A lower sensitivity and specificity will decrease the estimated

VE obtained from the evaluation. However, this might not be practical in settings relying on rapid tests. In some instances, an algorithm can be applied to decrease the chance of misclassification. For example, in cases where specificity is very high, but sensitivity is less than ideal, testing can first start with a rapid test, and those who test positive are considered true positive and do not undergo further testing. Those who test negative might be a false negative and should undergo another round of testing with rRT-PCR, with the rRT-PCR results being considered the test results for classification as a case or not. Additionally, statistical methods might be able to be applied to correct for imperfect sensitivity and specificity; consultation with a statistician prior to launching a VE evaluation is recommended. Antibody testing should not be used as the primary method to classify participants as cases or controls, due to unknown timing of infection and false-negative results.

8.1.2 Specimen collection

While analytic sensitivity and specificity of rRT-PCR is high, clinical sensitivity and specificity could be much lower for a variety of reasons, including poor specimen quality, specimen collection late in the course of disease, or improper specimen handling *(78)*. Furthermore, due to the need to enrol patients who will likely have detectable SARS-CoV-2, laboratory-confirmed outcomes should be restricted to patients who have a specimen collected within 10 days of illness onset.

8.1.3 Specimen type

SARS-CoV-2 can be detected in several body fluids and compartments, and the optimal specimen depends on clinical presentation and time since symptom onset *(81)*. At minimum, respiratory specimens should be collected, as the virus is most frequently detected in respiratory material *(78)*.

- **Upper respiratory tract (URT) specimens** are adequate for testing early stage infections, especially in asymptomatic or mild cases. Testing combined nasopharyngeal and oropharyngeal swabs from one individual has been shown to increase sensitivity for detection of respiratory viruses and improve the reliability of the result. A few studies have found that individual nasopharyngeal swabs yield a more reliable result than oropharyngeal swabs. Saliva and nasal swabs are not recommended as a diagnostic specimens for SARS-CoV-2 by WHO as this time *(78)*.

- **Lower respiratory tract (LRT) specimens** are advised if collected later in the course of the COVID-19 disease or in patients with a negative URT sampling and there is a strong clinical suspicion of COVID-19. LRT specimens can consist of sputum, if spontaneously produced (induced sputum is not recommended as this poses an increased risk of aerosol transmission) and/or endotracheal aspirate or bronchoalveolar lavage in patients with more severe respiratory disease, though appropriate precautions must be taken as these are aerosol-generating procedures.

For the purposes of a VE evaluation, it is recommended that all persons meeting the enrolment case definition have a nasopharyngeal and/or an oropharyngeal swab specimen to determine their case status.

8.2 Genomic characterization

SARS-CoV-2 has developed multiple mutations, leading some viruses to be deemed variants of concern 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 *(82)*. Genomic characterization within the context of VE evaluations offers an opportunity to contribute to the global understanding of variants of concern, including whether current vaccines protect against them.

If possible, all cases in VE evaluations, regardless of vaccine status, should have their specimen undergo sequencing. If a sufficient number of cases have their samples sequenced, one could determine the VE against commonly circulating variants. In the setting of multiple variants in circulation, sample size will likely need to be inflated to estimate variant-specific VE. Influenza VE studies routinely evaluate VE against influenza type/subtype separately, and in some analyses by clade *(7, 83)*. If a country is unable to perform sequencing, samples could be sent, along with information about vaccination, to a SARS-CoV-2 reference laboratory for genomic characterization *(84)*. All sequence data should be shared, with relevant vaccination meta-data, through publicly available databases *(78)*. Of note, if sequencing of all cases is not possible, then sequencing a subset of cases to document the circulating variant composition of the population will allow the results to be interpreted in the context of the predominant variants in circulation.

8.3 Testing for prior infection

Some enrolled participants in a VE evaluation will have had prior infection, potentially providing them with natural immunity. This can bias the VE estimate, as described previously in Section 7. Bias in VE studies of COVID-19 vaccines. We recommend ascertaining prior disease or infection based on history. Additional accuracy in identifying prior infection could be achieved by collecting blood samples and using serological testing for antibodies against SARS-CoV-2 in cohort studies or among controls, though this will increase the complexity and cost of the evaluation and may not be feasible in most settings. The performance of serologic assays varies widely in different testing groups (such as in patients with mild versus moderate-to-severe disease as well as in young versus old), timing of testing and the target viral protein *(78)*. Furthermore, antibody detection tests for SARS-CoV-2 may also rarely cross-react with common coronaviruses or other pathogens and thus yield false-positive results. Antibody testing for the purpose of evaluation of prior history of COVID-19 can be conducted with lateral flow immunoassays (LFI), enzyme-linked immunosorbent assays (ELISA), or chemiluminescence immunoassays (CLIA). Careful consideration should be made as to which antibody target is selected, as some targets, such as the antibody against the spike protein could be generated from both vaccination and disease, and N protein antibodies tend to wane faster after infection.

9. Statistical considerations

9.1 Sample size

As with any epidemiologic study, COVID-19 VE evaluations must be large enough to rule out chance as an explanation for the evaluation results. The sample size needed to estimate VE will depend on a number of factors, including the proportion of the population that is vaccinated, the incidence of the evaluation outcome, the expected VE, and the desired precision of the VE estimates. Sample sizes can be calculated following the methodology described by O'Neill *(85)*. The formulas calculate a minimum sample sizes and will need to be inflated to account for nonparticipation, stratification, adjustment for effect modifiers and confounding, or exclusion factors that might be identified only after enrolment. Note, VE should be calculated for each vaccine regimen separately (e.g. product A dose 1, product A dose 2 is one regimen while product A dose 1, product B dose 2 is a different regimen). Thus, if participants in the evaluation are expected to be vaccinated with a mix of vaccine products, the sample size will need to be increased to ensure sufficient persons are enrolled based on the population coverage of each regimen. For the below calculations, we have selected a precision of ± 10%, but this can be adjusted based on local needs and resources. While a wider precision interval will result in a lower required sample size it will also lead to less certainty in the interpretation of the VE estimate and will make secondary analyses that use a subset of the sample more challenging.

9.1.1 Cohort study

For a cohort study, a formula *(85)* to calculate the minimum sample size is:

$$N = (z/d)^2(1/ARU)((1-ARU)/ARU)+(1-ARU)/ARU)$$
$$= (z/d)^2((1+1/\Psi)/ARU-2)$$

Where z denotes the (1-α) percentage point of the standardized normal distribution (normally this is based on an α = 0.05 and thus z = 1.96); ARU denotes the assumed attack rate in the unvaccinated group, ψ = 1-VE, where VE denotes the anticipated vaccine effectiveness; and d is determined by solving the equation $W(\hat{\beta},\hat{d}) = \exp(\hat{\beta})(\exp(\hat{d})-(\exp(-\hat{d}))$ where $\hat{d}=z\hat{\sigma}$ where $W(\hat{\beta},d)$ denotes the confidence interval (CI) width, i.e. the difference between the upper and lower limits.

Table 4 gives the minimum number of persons to enrol in a cohort study to detect the specified VE against COVID-19 disease outcomes, assuming different attack rates in the unvaccinated, and a precision of ± 10%, with a type 1 error rate (α) of 0.05. Achieving greater precision than ± 10% leads to large sample size requirements in a cohort study and is unlikely to be feasible; longer follow-up times might reduce the sample size required. Furthermore, as the attack rate among the unvaccinated increases (i.e. high transmission in community) the sample size needed decreases. If the vaccine is highly effective, and coverage is high, then the attack rate might decrease among both unvaccinated and vaccinated due to herd protection, requiring larger sample sizes than when coverage is lower and there is more transmission.

Table 4. Minimum number of persons to enrol in a cohort study to detect the specified VE, assuming different attack rates in the unvaccinated, and a precision of ± 10%, with a type 1 error rate of 0.05

Vaccine effectiveness	Attack rate unvaccinated		
	1%	2%	5%
50%	28 998	14 402	5644
60%	21 830	10 852	4266
70%	15 454	7691	3034
80%	9921	4944	1958
90%	5430	2710	1079

9.1.2 Case-control and test-negative design study

For a case-control or test-negative design study, the formula *(85)* to calculate the minimum sample size of cases (N_1) is:

$$N_1 = (z/d)^2[1/A(1-A)+1/CP_2(1-P_2)]$$

Where C is the control to case ratio (e.g. C = 2 denoting 2 controls for every case); P_2 denotes the prevalence of vaccine exposure in the control group (i.e. vaccine coverage in the population being studied); $A = P_2(1-VE)/[1-P_2(VE)]$ where VE denotes the anticipated vaccine effectiveness; z denotes the (1-α) percentage point of the standardized normal distribution (normally this is based on an α = 0.05 and thus z = 1.96); and d is determined by solving the equation $W(\hat{\beta},\hat{d}) = \exp(\hat{\beta})(\exp(\hat{d})-(\exp(-\hat{d}))$ where $\hat{d}=z\hat{\sigma}$ where $W(\hat{\beta},d)$ denotes the CI width, i.e. the difference between the upper and lower limits. The number of controls needed are then calculate as C*N1.

Annex 2: Sample size (👆) gives the minimum number of persons to enrol in a case-control or test-negative design study to detect the specified VE, assuming different vaccine coverages ranging from 20%–90%, with 1–4 controls per case, and a precision of ± 5% and ± 10%, with a type 1 error rate of 0.05. Table 5 is a brief synopsis to highlight the sample sizes needed. It should be noted that for the most part, higher VE results in smaller samples sizes as does higher coverage in the control group. However, sample size again increases when coverage is greater than 75%.

Table 5. Minimum number of cases and controls to detect for a specified VE, estimated vaccination coverage in the population under evaluation, with 1–3 controls per case, with a precision of ± 10%, and a type 1 error rate of 0.05

Vaccine effectiveness	Vaccination coverage in the population being studied	1:1 cases to controls		1:2 cases to controls		1:3 cases to controls	
		No. cases	No. controls	No. cases	No. controls	No. cases	No. controls
50%	30%	1133	1133	902	1804	825	2475
	50%	828	828	633	1266	568	1704
	70%	855	855	624	1248	546	1638
	90%	1736	1736	1195	2390	1015	3045
70%	30%	526	526	441	882	412	1236
	50%	346	346	274	548	250	750
	70%	319	319	234	468	205	615
	90%	580	580	381	762	315	945
90%	30%	150	150	138	276	134	402
	50%	80	80	70	140	67	201
	70%	56	56	45	90	41	123
	90%	75	75	48	96	39	117

For each case, 1 to 4 control subjects could be identified. Enrolling more than 4 controls per case does not increase power substantially. Increasing the number of controls per case helps to increase the efficiency, thus decreasing the number of cases that need to be identified and can be useful in settings where incidence of COVID-19 is low or in smaller sentinel sites which might not be able to enrol enough cases. In many settings, where numbers of COVID-19 cases are large, it will likely be more efficient to increase the number of cases enrolled rather than the number of controls per case.

For case-control studies, the method of sample size calculation proposed here pertains to unmatched designs. The sample size needed for matched case-control designs requires methodology that accounts for the matching factors and is beyond the scope of this document *(85)*.

9.2 Matching

Selecting controls to be similar to case patients in particular characteristics in case-control studies, or selecting unexposed and exposed individuals to be similar in cohort studies is termed "matching". The purpose of matching in cohort studies is to efficiently diminish the confounding effects of various factors, particularly those that are difficult to adjust for in the analysis or factors that can be more efficiently addressed by the study design. In case-control studies, matching can be performed to reduce variance and improve precision but will create bias unless the matching is accounted for in the analysis *(86, 87)*. Matching presents challenges with the loss of flexibility in enrolment and potential for introduction of a bias and imprecision due to overmatching. Matching on age to improve efficiency could be considered, depending on the study setting and recommended vaccine target groups, as it is likely that VE among certain age groups will likely be of public health interest. This can be achieved by frequency matching, potentially in 10-year age bands if sufficient cases and non-cases/controls in each age band are enrolled. It would be ideal to be able to match on other COVID-19 risk factors, but this is challenging to do practically given the inability to quantify these risk factors easily. If conducting a multisite evaluation, it is suggested to frequency match on enrolment site (or neighbourhood) to help account for differences between populations at the different sites, such as sociodemographic factors that might be associated with vaccine and disease status. For TND and case-control studies, participants should be matched by time of enrolment (or this should be adjusted for in the analysis), as epidemiology of SARS-CoV-2 is constantly changing as are the circulating variants and controls should be selected at a similar time period as cases.

9.3 Data collection, management and analyses

9.3.1 Data collection and management

Data should be collected to characterize study setting, including 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. Individual-level data should be collected by interviewing participants using a study questionnaire, medical record review, laboratory record review, and vaccine history review. An accurate, detailed vaccination history, including the vaccine product administered and dates of vaccination, is critical and should ideally be obtained from written records (see Section 4. Assessing COVID-19 vaccination history ( )). Data-collection forms should not include any identifiable information (e.g. name) but instead use unique identifiers. A separate form should be maintained that links the identifiers with participant names, and confidentiality must be maintained. Data collection should comply with local data protection requirements. Data from each evaluation site should be entered into the database and reviewed for completeness and any data entry errors. The database should undergo a quality control process.

9.3.2 Characterizing participants

Several descriptive analyses are useful for characterizing the participants. Graphs of case counts over time are informative for understanding the COVID-19 epidemic in the target population. These plots can

help in assessing whether calendar time is a potential confounder, particularly if coupled with plots of vaccine uptake over time. Flowcharts of the enrolment process can be used to identify possible sources of bias such as high rates of refusal to participate in the evaluation.

For the primary VE analysis, fully vaccinated participants should be compared with unvaccinated subjects. Additionally, persons with prior infection should be included in the primary analysis in settings where vaccination is given to all individuals in target groups regardless of prior infection status. In settings, where prior infection is an exclusion from vaccination, the primary analysis should also be limited to those without known prior infection.

Bivariate descriptive statistics are used to assess the distribution of covariates among the participants. These are particularly useful for comparing covariate distributions between vaccinated versus unvaccinated subjects and between COVID-19 cases versus non-cases/controls. These bivariate comparisons aid in determining variables that may be included as potential confounders in adjusted VE estimates or variables by which to stratify the analyses. The bivariate distributions are also useful for comparing the participants in your evaluation with participants in other COVID-19 VE evaluations, as differences in evaluation populations may contribute to different VE estimates.

9.3.3 Analytic approaches to estimate crude rate/odds ratios

In cohort studies, unadjusted (or crude) rate ratios can be obtained from the ratio of incidence rates of COVID-19 outcomes between vaccinated and unvaccinated subjects. Unadjusted incidence rate ratios can be calculated using 2×2 tables or regression models (e.g. Poisson, Cox-proportional hazards). In case-control and TND studies, unadjusted odds ratio estimates can be obtained from the ratio of the odds of exposure (i.e. vaccination) among controls compared with cases. If matching was conducted, matching should be factored into the analysis.

9.3.4 Assessing and adjusting for confounders

In planning the analyses, investigators should determine which measured covariates are acting as confounders (see Section 5. Measuring covariates ( )), by identifying variables that are associated with both vaccination and COVID-19. For this, the most appropriate method for the final selection of potential confounders to include in statistical models is the "change-in-estimate" approach *(88, 89)*. In the change-in-estimate approach, the unadjusted VE is estimated using an appropriate statistical model. VE is then estimated adjusting for a single covariate. If the adjusted VE differs from the unadjusted VE by more than a pre-determined percentage, the covariate is considered to be a confounder and will be included in final models. A common threshold is to include covariates whose adjustment changes the crude odds ratio/rate ratio by 10% or more, but the threshold is at the investigators' discretion.

9.3.5 Time since vaccination

COVID-19 VE evaluations differ from influenza VE evaluations because it is expected that people initially unvaccinated may soon become vaccinated within a targeted group due to the rapid rollout of vaccines. Moreover, unlike influenza vaccine that is usually given prior to the influenza season, COVID-19 vaccines can be given during periods of high incidence. Therefore, some people will likely be infected soon after vaccination, and so it will be especially important to define the time after vaccination when a person would be considered immunologically protected. Moreover, it is possible that some people who have systemic reactions to vaccination will present for COVID-19 testing in the days after vaccination, which would lead to a bias in the VE estimate, particularly in the TND.

Thus, for the primary analysis of VE, 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. However, while 14 days for the first dose and 7-14 days for the second dose has been used in clinical trials of the COVID-19 vaccines to date and is used most often for other vaccines, the number of days post-vaccination to use in the primary analysis should be driven by the specific product being evaluated (90). For cohort studies, this would be 14 days after vaccination and for case-control and TND studies, this would be 14 days prior to disease onset (for cases in a traditional case-control study and cases and controls in a TND), or hospitalization date (for hospitalized controls in a traditional case-control study). While the 14-day cut-off for considering a person protected from vaccination might reduce precision by excluding a certain number of cases (and vaccinated controls), it will optimize the validity of the VE estimate. We recommend doing secondary analyses that use shorter intervals of time after vaccination in considering people vaccinated (e.g. 7 days, 10 days); additional analyses can be done for discrete intervals of days in the 14-day period after vaccination to assess the onset of VE (e.g. 7–13 days). However, we recommend that investigators should not consider a person as unvaccinated in the few days after vaccination as the exact number of days to start to achieve protection is not known; this will minimize exposure misclassification bias (see Section 7. Bias in VE studies of COVID-19 vaccines ( )).

9.3.6 Final analyses of VE

Multivariable regression analysis allows adjustment for confounding variables. After identifying potential confounders, final VE estimates can be calculated using the adjusted odds ratio (aOR) or rate ratio (aRR) for vaccination using the following formula: VE = (1-(aOR or aRR)) × 100%. Regression modelling can also quantify and measure the precision of any effect modification, in which VE differs by subgroup. If effect modification exists, VE and CIs should be reported for each subgroup separately. VE should be calculated separately for each vaccine regimen and for each product. At this time there is no WHO recommendation to support mixing vaccine types for first and second doses, although the VE for mixed vaccine schedules will be of interest if this has occurred in the study population and, if sample size allows, should be evaluated. As with any analyses, appropriateness of the statistical methods should be reviewed with a statistician with respect to model diagnostics and validity checks (e.g. goodness of fit, identification of outliers, and assessment for multiple co-linearity).

9.3.7 Additional analyses

For the primary VE analysis, fully vaccinated participants should be compared with unvaccinated subjects. For a secondary analysis, partially vaccinated subjects could be compared with unvaccinated subjects to determine if partial vaccination is effective. Stratified analyses, based on subgroups and effect modifiers, as well as common viral variants in circulation, are important to do if sufficient sample size allows acceptable precision. Another secondary analysis of importance in areas with incomplete documentation of vaccination status would be to accept verbal reports in addition to documented reports of vaccination. Lastly, a secondary analysis can be done that excludes persons with a history of prior COVID-19, potentially stratifying by lab-confirmed prior infection and clinically diagnosed and/or stratifying by the time since prior infection if the sample size is large enough.

Additionally, a sensitivity analysis should be considered for some key variables to assess the robustness of the VE estimate. Much is unknown still about COVID-19, and a sensitivity analysis allows one to understand the impact of various choices for inclusion/exclusion into the final VE calculation. Examples of potential sensitivity analyses include:

- A sensitivity analysis of importance in areas with incomplete documentation of vaccination status would be to accept verbal reports in addition to documented reports of vaccination, and compare this

with an analysis whereby all with verbal report are considered unvaccinated or vaccinated to assess how the extremes of exposure misclassification would impact the VE.

- A sensitivity analysis with respect to timing of the test in relationship to symptom onset. As cases further into their clinical course might be less likely to be positive for SARS-CoV-2, VE can be evaluated among those testing early as well as those testing later in relation to the course of illness to see the impact of these potentially false negatives *(81)*.

If a sufficient number of breakthrough cases are identified as part of the VE evaluation, then investigators can consider doing an analysis looking at risk factors for breakthrough. However, most VE evaluations will not be powered or designed to be able to assess risk factors for breakthrough infections adequately.

9.3.8 Interpretation and extrapolation of results from VE evaluations

The efficacy of COVID-19 vaccinations will have been established by a limited number of pre-authorization clinical trials; the findings of any VE evaluation should therefore be interpreted in light of the pre-authorization trials. If VE is found to be different than expected, it is particularly important that further investigation should be conducted (Box 4), including an examination of the vaccine management and vaccine administration techniques. The results can then be used to take corrective action, if necessary *(65)*. The implementation of the evaluation and methods for analysis should also be examined to ensure that case definitions were applied consistently, that case ascertainment was appropriate, vaccination status was appropriately determined, that known confounders were controlled for and that identifiable biases did not occur.

Unexpected findings should be followed up with a detailed programmatic and epidemiological evaluation. It is important to note that a VE that is lower than expected, or even negative, should not be immediately interpreted as evidence of vaccine-associated enhanced disease (VAED); more common causes, such as bias, population differences, or waning protection, should first be considered. Moreover, confirmation of VAED would likely need further investigation, including biomarker evaluations *(91)*.

BOX 4. POTENTIAL REASONS FOR VE ESTIMATES THAT ARE DIFFERENT FROM VACCINE EFFICACY RESULTS

VE estimate valid	VE estimate not valid
• 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 VAED (especially severe disease outcome) • Prevalence of prior infection in population different from that of efficacy study	• Error in implementation (e.g. enrolment of persons not meeting case definition, poor specimen collection/handling) • Biases • Unmeasured or incompletely controlled confounders • Chance finding; more likely with small sample size

9.3.9 Pooling data from multiple VE evaluations

As noted above, robust sample sizes are needed to estimate VE, which can limit the assessment of VE in important subgroups. One approach to enable stratified estimates by subgroups is to pool data from separate VE evaluations. This can be done through meta-analysis of reported VE estimates, or through the pooling of individual-level subject data. In either case, pooling data has several challenges that must be considered. The evaluations being pooled need to be measuring the same outcome, for the same vaccine product, at the same type of site of case capture (e.g. hospital vs outpatient clinic). Pooled evaluations must be sufficiently similar in terms of case definitions, exclusion criteria, and vaccine status definitions. The evaluations to be pooled must have similar data available on key covariates to include in adjusted VE models. Perhaps most importantly, the evaluation settings must be similar enough that the pooled results can be generalized. This means that the populations under evaluation must have comparable access to vaccination and to health care for COVID-19 illness. For these reasons, **we caution against pooling data from populations that are heterogeneous with respect to the following: vaccine programmes or policies, health systems or care seeking behaviours, or infection risk overall.** This would apply to pooling data from special populations, such as those in prisons or nursing homes, with general community-dwelling populations. If pooled analyses are to be attempted, qualitative and statistical heterogeneity between populations should be assessed, and the best methods for meta-analysis or pooling data should be determined in consultation with statisticians *(92–94)*. Pooling individual-level data from multisite or multicountry studies could be done if the same protocol is used a priori, as has been done in multicountry evaluations of influenza VE, but there is still a need to account for heterogeneity *(95, 96)*.

10. Platforms to do COVID-19 VE evaluations

A potentially efficient approach to conducting COVID-19 evaluations is to build an evaluation onto an existing surveillance platform used for another purpose. However, various factors need to be considered when deciding if an existing platform could be leveraged to conduct a COVID-19 VE evaluation.

Some factors to consider when deciding if an existing platform can be used include the following:

- ☐ The ability to achieve the evaluation objective by using the existing platform (e.g. does it capture the outcomes of interest?).
- ☐ The alignment of the suspected COVID-19 case definition with the case definition used by the existing platform, including inclusion and exclusion criteria.
- ☐ The population being studied by the existing platform is eligible for vaccination.
- ☐ The ability to recruit the target sample size.
- ☐ The modifications to the case investigation forms that need to be made to ensure that COVID-19 vaccination history, potential COVID-19 confounders, and COVID-19 risk factors are collected.
- ☐ The specimens currently collected and the ease of adding appropriate COVID-19 specimen collection onto the platform.
- ☐ The ability to conduct high-quality COVID-19 testing using existing laboratory capacity.
- ☐ The additional resources that are needed to build a COVID-19 VE evaluation onto the existing platform (e.g. collecting documentation of vaccination).

Some of the following platforms are already in place in countries and could be modified to conduct a VE evaluation.

- ☐ **SARI surveillance:** SARI surveillance is usually undertaken to identify circulating influenza strains and to conduct influenza VE evaluations. The WHO recommended SARI case definition is a hospitalized 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 *(39)*. It is used to identify suspected cases for enrolment and can be used without modification given that it can capture COVID-19 cases, but it is not so specific as to exclude persons without COVID-19, making it amenable to TND studies. Of note, the SARI case definition includes fever, and not all COVID-19 patients have fever, so some cases might be missed, which could introduce a bias if COVID-19 symptoms are modified by vaccination. In addition, the same specimen, usually a nasopharyngeal swab, is used to test for both influenza and COVID-19. The SARI platform is best suited to understand VE for severe outcomes as all patients are hospitalized. Examples of SARI networks include the Global Influenza Surveillance and Response System (GISRS); Red para la Evaluación de la Efectividad de la Vacuna en Latino América y el Caribe – influenza (REVELAC-i); Influenza - Monitoring Vaccine Effectiveness in Europe (I-MOVE) network; and the Global Influenza Hospital Surveillance Network *(97–100)*.

- **ILI surveillance:** ILI surveillance is undertaken for similar reasons as SARI surveillance, but usually captures patients at outpatient facilities. Therefore, the ILI platform is best suited to understand VE for non-severe disease, though the caveat is that an unknown proportion of ILI cases might go on to hospitalization/severe disease. An ILI case is defined as a person with an acute respiratory infection with measured fever of ≥ 38 C°, cough, with onset within the last 10 days *(39)*. As with the SARI definition, the ILI definition also includes fever so will not capture all COVID-19 cases. Some primary care networks, such as I-MOVE, already participate in ILI surveillance and could be adapted to conduct COVID-19 VE evaluations *(97)*.

- **Inpatient sentinel surveillance for other diseases:** Some countries have inpatient sentinel site surveillance to identify outbreaks, to understand disease burden of known pathogens, and/or to identify novel pathogens. Examples of sentinel site surveillance includes those for acute febrile illness, pneumonia/LRT infection, invasive bacterial disease and acute gastroenteritis. Existing sentinel surveillance platforms can offer a field team familiar with enrolment following a standard case definition, specimen collection, and access to reliable clinical laboratories. However, unlike SARI and ILI platforms, URT specimens might not be collected as part of sentinel surveillance platforms, and so would require additional specimen collection. Moreover, case definitions for other diseases targeted by sentinel surveillance are unlikely to capture sufficient numbers of COVID-19 cases, and if unmodified might capture atypical clinical presentations (e.g. diarrhoea, fever without respiratory symptoms), which could lead to biased VE estimates.

- **Health worker surveillance or cohorts:** In some countries, health worker surveillance or cohort studies are being undertaken to identify risk factors for infection, determine serial seroprevalence, or as part of an institutional infection control and prevention strategy. Such surveillance among health workers can be amenable to VE studies, particularly cohort designs. As part of the Unity Studies, WHO has developed a health worker cohort study that could be modified to conduct VE studies *(101)*.

- **COVID-19 adverse events following immunization (AEFI) studies:** AEFI studies might be able to be leveraged to evaluate duration of immunity and in some cases, to estimate VE. The advantage of using an AEFI platform is that detailed information on vaccination history will already be collected. Moreover, for most COVID-19 AEFI studies, severe COVID-19 disease will likely be an outcome of interest due to the concern of VAED. In places where a cohort of only vaccinated persons are followed (cohort event monitoring [CEM]), one can follow the cohort to determine if and when persons develop COVID-19 disease after vaccination, and an analysis can be done to understand the duration of vaccine protection *(102)*. However, to evaluate COVID-19 VE, a comparison group of unvaccinated persons would need to be added to AEFI CEM platforms. Additionally, most AEFI CEM studies end after 3 months, and a longer duration (at least 1 year) of follow-up would be valuable to understand the duration of protection.

- **Administrative databases:** In some settings, detailed medical records or administrative databases that document both disease outcomes and vaccination history might exist or could be created through record linkage. A single comprehensive database for participants in managed health care organizations or large hospitals might be available, making extraction of key variables for both COVID-19 cases, and a set of controls, efficient. In other settings, separate databases might be linked by patient identification numbers to extract all the requisite variables. However, such comprehensive databases are unlikely to exist in most L/MICs. When evaluating these databases, it is important to understand potential testing algorithms in place and the sensitivity/specificity of each variable. Each database needs to be evaluated to ensure quality.

- **Outbreaks:** Outbreaks of COVID-19 can serve as settings to undertake VE evaluations efficiently. Outbreaks in well-defined populations, such as long-term care facilities, military barracks, prisons, hospitals or schools, offer ideal settings. Statistical efficiency is greatest when a proportion of the outbreak population, approximately 30%–70%, has been vaccinated at least 2 weeks before the onset of the outbreak as part of routine rollout of vaccines. Other characteristics of the outbreak that optimize VE evaluations are to institute COVID-19 testing of all persons during the outbreak period, or at least all symptomatic persons, as well as enough cases to have good precision around the VE estimate (e.g. > 30 cases). Both cohort and case-control studies can be undertaken in outbreak settings. As mentioned, the screening method might also be applicable as the vaccine coverage among cases and the entire outbreak population will likely be known; the screening method can be particularly useful in anticipating the expected number of vaccinated cases *(65)*. Moreover, if outbreaks occur in multiple facilities, some of which have been vaccinated and some not, comparison of the outbreak size, duration and severity can be made to assess the vaccine's impact on outbreaks. Of note, COVID-19 vaccination as a response measure to an outbreak is not being widely recommended at this time due to time lag in achieving population immunity. VE evaluations in outbreak settings are susceptible to the same biases as all VE evaluations, including some additional ones such as the implementation of non-pharmaceutical and pharmaceutical (e.g. monoclonal antibodies) measures as part of the outbreak response.

11. Protection of human subjects and informed consent

COVID-19 vaccines introduced into countries will have been authorized by national regulatory authorities as safe for routine use. Participants in VE studies will not have been randomized to receive vaccine and will likely be identified through routine clinical testing and/or public health surveillance. Moreover, demonstrating COVID-19 VE through studies in the groups targeted for vaccination will provide essential information to ministries of health to assess the health and public health benefits of the vaccination programme, and inform future policy decisions. For these reasons, in some countries, VE evaluations might be deemed public health programme evaluation and/or surveillance and be given non-research determination. In many countries, this determination is done by an expedited review of the protocol by an ethical review committee (ERC). The ERC might determine to exempt the protocol as non-research, and therefore not require informed consent, or they might require full ethical review and approval. In the latter case, the ERC could determine that either verbal or written informed consent is required from participants. Specific consent procedures may be needed for unconscious or critically ill patients who are unable to give written consent of their own volition (e.g. oral witnessed consent, consent by the next of kin). Informed consent should outline risks and benefits of participation in the local language. Risks to participants might be deemed those incurred by additional specimen collection (apart from those done for routine clinical care), or if any questions on the questionnaire are felt to be of a sensitive nature. Benefits to participants might be access to COVID-19 vaccines if they have not yet been vaccinated. The voluntary nature of participation and the ability to withdraw consent without fear of reprisals should be clearly stated. If consent is required, parents should be asked to provide it for their child, supplemented by assent by the child (depending on their age). Templates for informed consent are available on the WHO website *(103)*. Regardless of ERC determination, all data must be protected, and confidentiality of participants must be ensured.

12. Reporting of results

WHO encourages consistent and standardized reporting of results of COVID-19 VE evaluations for several reasons. As with all observational studies, reports of COVID-19 VE evaluations should 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 done 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 *(104, 105)*. Lastly, as VE studies of COVID-19 vaccines start becoming available, having a standardized format for reporting will facilitate ease of interpretation for the many audiences that will be interested in these 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 *(106)*. 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 enrolment 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 specific features of VE studies and unique aspects of COVID-19 epidemiology and vaccines, additional data elements for COVID-19-specific VE studies expanding upon the STROBE checklist are recommended. We recommend reporting on additional elements, as outlined in Annex 4: Reporting (). Standardized reporting of observational studies of influenza VE studies, as well as other vaccines, by adapting STROBE recommendations has been recommended by WHO's Immunization and Vaccine-related Implementation Research Advisory Committee (IVIR-AC) *(107)*. Such adaptations of STROBE guidelines to specific categories of observational studies have been done before *(108)*. While many biomedical journals have word count limits which may restrict reporting of all of these elements, most allow for online supplements which can accommodate the additional data elements. Although not part of the STROBE guidelines, WHO encourages sharing of COVID-19 VE evaluation databases in data repositories available to the public, to encourage transparency and facilitate pooling of results *(109)*.

References

1. WHO Coronavirus Disease (COVID-19) Dashboard. Geneva: World Health Organization; 2021 (https://covid19.who.int/, accessed 2 March 2021).
2. Draft landscape and tracker of COVID-19 candidate vaccines. Geneva: World Health Organization; 2021 (https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines, accessed 2 March 2021).
3. Beaumont P. Covid-19 vaccine: who are countries prioritising for first doses? The Guardian. 18 November 2020 (https://www.theguardian.com/world/2020/nov/18/covid-19-vaccine-who-are-countries-prioritising-for-first-doses, accessed 2 March 2021).
4. WHO SAGE values framework for the allocation and prioritization of COVID-19 vaccination. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/handle/10665/334299, accessed 2 March 2021).
5. Measuring impact of *Streptococcus pneumoniae* and *Haemophilus influenzae* type b conjugate vaccination. Geneva: World Health Organization; 2012; WHO/IVB/12 (https://apps.who.int/iris/handle/10665/75835, accessed 2 March 2021).
6. Generic protocol for monitoring impact of rotavirus vaccination on gastroenteritis disease burden and viral strains. Geneva: World Health Organization; 2008; WHO/IVB/08 (https://apps.who.int/iris/handle/10665/69913, accessed 2 March 2021).
7. Evaluation of influenza vaccine effectiveness. Geneva: World Health Organization; 2017 (https://apps.who.int/iris/handle/10665/255203, accessed 2 March 2021).
8. Roadmap to improve and ensure good indoor ventilation in the context of COVID-19. Geneva: World Health Organization; 2021 (https://apps.who.int/iris/rest/bitstreams/1333991/retrieve, accessed 2 March 2021).
9. Mask use in the context of COVID-19: interim guidance. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/rest/bitstreams/1319378/retrieve, accessed 2 March 2021).
10. Tong ZD, Tang A, Li KF, Li P, Wang HL, Yi JP et al. Potential presymptomatic transmission of SARS-CoV-2, Zhejiang Province, China, 2020. Emerg Infect Dis. 2020;26:1052-4. doi: 10.3201/eid2605.200198.
11. Wei WE, Li Z, Chiew CJ, Yong SE, Toh MP, Lee VJ. Presymptomatic transmission of SARS-CoV-2 - Singapore, January 23-March 16, 2020. MMWR Morb Mortal Wkly Rep. 2020;69:411-5. doi: 10.15585/mmwr.mm6914e1.
12. Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR et al. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. N Engl J Med. 2020;382:2081-90. doi: 10.1056/NEJMoa2008457.
13. Li J, Huang DQ, Zou B, Yang H, Hui WZ, Rui F et al. Epidemiology of COVID-19: a systematic review and meta-analysis of clinical characteristics, risk factors, and outcomes. J Med Virol. 2020. doi: 10.1002/jmv.26424.
14. Yang J, Zheng Y, Gou X, Pu K, Chen Z, Guo Q et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. Int J Infect Dis. 2020;94:91-5. doi: 10.1016/j.ijid.2020.03.017.
15. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. BMJ. 2020;369:m1985. doi: 10.1136/bmj.m1985.
16. Lu L, Zhong W, Bian Z, Li Z, Zhang K, Liang B et al. A comparison of mortality-related risk factors of COVID-19, SARS, and MERS: a systematic review and meta-analysis. J Infect. 2020;81:e18-e25. doi: 10.1016/j.jinf.2020.07.002.
17. Noor FM, Islam MM. Prevalence and associated risk factors of mortality among COVID-19 patients: a meta-analysis. J Community Health. 2020;45:1270-82. doi: 10.1007/s10900-020-00920-x.
18. Kumar A, Arora A, Sharma P, Anikhindi SA, Bansal N, Singla V et al. Clinical features of COVID-19 and factors associated with severe clinical course: a systematic review and meta-analysis. SSRN. 2020:3566166. doi: 10.2139/ssrn.3566166.
19. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395:497-506. doi: 10.1016/S0140-6736(20)30183-5.
20. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. BMJ. 2020;369:m1966. doi: 10.1136/bmj.m1966.
21. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395:1054-62. doi: 10.1016/S0140-6736(20)30566-3.

22. Tian W, Jiang W, Yao J, Nicholson CJ, Li RH, Sigurslid HH et al. Predictors of mortality in hospitalized COVID-19 patients: a systematic review and meta-analysis. J Med Virol. 2020;92:1875-83. doi: 10.1002/jmv.26050.

23. Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and mortality among black patients and white patients with Covid-19. N Engl J Med. 2020;382:2534-43. doi: 10.1056/NEJMsa2011686.

24. Gold JAW, Wong KK, Szablewski CM, Patel PR, Rossow J, da Silva J et al. Characteristics and clinical outcomes of adult patients hospitalized with COVID-19 - Georgia, March 2020. MMWR Morb Mortal Wkly Rep. 2020;69:545-50. doi: 10.15585/mmwr.mm6918e1.

25. Palaiodimos L, Kokkinidis DG, Li W, Karamanis D, Ognibene J, Arora S et al. Severe obesity, increasing age and male sex are independently associated with worse in-hospital outcomes, and higher in-hospital mortality, in a cohort of patients with COVID-19 in the Bronx, New York. Metabolism. 2020;108:154262. doi: 10.1016/j.metabol.2020.154262.

26. Zambrano LD, Ellington S, Strid P, Galang RR, Oduyebo T, Tong VT et al. Update: characteristics of symptomatic women of reproductive age with laboratory-confirmed SARS-CoV-2 infection by pregnancy status - United States, January 22-October 3, 2020. MMWR Morb Mortal Wkly Rep. 2020;69:1641-7. doi: 10.15585/mmwr.mm6944e3.

27. Multisystem inflammatory syndrome in children and adolescents with COVID-19. Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/multisystem-inflammatory-syndrome-in-children-and-adolescents-with-covid-19, accessed 2 March 2021).

28. Mullard A. How COVID vaccines are being divvied up around the world. Nature. 30 November 2020 (https://www.nature.com/articles/d41586-020-03370-6, accessed 2 March 2021).

29. COVAX Facility [website]. Gavi, the Vaccine Alliance; 2021 (https://www.gavi.org/covax-facility, accessed 2 March 2021).

30. Fair allocation mechanism for COVID-19 vaccines through the COVAX Facility. Geneva: World Health Organization; 2020 (https://www.who.int/publications/m/item/fair-allocation-mechanism-for-covid-19-vaccines-through-the-covax-facility, accessed 2 March 2021).

31. Use of Emergency Use Listing procedure for vaccines against Covid-19 Q&A. Geneva: World Health Organization; 2020 (https://www.who.int/medicines/regulation/prequalification/prequal-vaccines/resources/QA_EUL_Covid-19_July2020.pdf?ua=1, accessed 2 March 2021).

32. Hanquet G, Valenciano M, Simondon F, Moren A. Vaccine effects and impact of vaccination programmes in post-licensure studies. Vaccine. 2013;31:5634-42. doi: 10.1016/j.vaccine.2013.07.006.

33. Halloran ME, Haber M, Longini IM, Jr., Struchiner CJ. Direct and indirect effects in vaccine efficacy and effectiveness. Am J Epidemiol. 1991;133:323-31. doi: 10.1093/oxfordjournals.aje.a115884.

34. Grijalva CG, Nuorti JP, Arbogast PG, Martin SW, Edwards KM, Griffin MR. Decline in pneumonia admissions after routine childhood immunisation with pneumococcal conjugate vaccine in the USA: a time-series analysis. Lancet. 2007;369:1179-86. doi: 10.1016/S0140-6736(07)60564-9.

35. Rossman H, Shilo S, Meir T, Gorfine M, Shalit U, Segal E. Patterns of COVID-19 pandemic dynamics following deployment of a broad national immunization program. medRxiv. 2021. doi:10.1101/2021.02.08.21251325 (https://www.medrxiv.org/content/10.1101/2021.02.08.21251325v1, accessed 2 March 2021).

36. Nordling L. Africa's pandemic puzzle: why so few cases and deaths? Science. 2020;369:756-7. doi: 10.1126/science.369.6505.756.

37. Aron J, Muellbauer J, Giattino C, Ritchie H. A pandemic primer on excess mortality statistics and their comparability across countries. Our World in Data; 2020 (https://ourworldindata.org/covid-excess-mortality, accessed 2 March 2021).

38. Clinical management of COVID-19: interim guidance. Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/clinical-management-of-covid-19, accessed 2 March 2021).

39. WHO surveillance case definitions for ILI and SARI. Geneva: World Health Organization; 2021 (https://www.who.int/influenza/surveillance_monitoring/ili_sari_surveillance_case_definition/en/, accessed 2 March 2021).

40. WHO COVID-19: case definitions. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/rest/bitstreams/1322790/retrieve, accessed 2 March 2021).

41. Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. Lancet. 2021;397:99-111. doi: 10.1016/S0140-6736(20)32661-1.

42. Household transmission investigation protocol for 2019-novel coronavirus (COVID-19) infection. Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/household-transmission-investigation-protocol-for-2019-novel-coronavirus-(2019-ncov)-infection, accessed 2 March 2021).

43. Preziosi MP, Halloran ME. Effects of pertussis vaccination on transmission: vaccine efficacy for infectiousness. Vaccine. 2003;21:1853-61. doi: 10.1016/s0264-410x(03)00007-0.

44. Crowcroft NS, Klein NP. A framework for research on vaccine effectiveness. Vaccine. 2018;36:7286-93. doi: 10.1016/j.vaccine.2018.04.016.

45. Cardemil CV, Dahl RM, James L, Wannemuehler K, Gary HE, Shah M et al. Effectiveness of a third dose of MMR vaccine for mumps outbreak control. N Engl J Med. 2017;377:947-56. doi: 10.1056/NEJMoa1703309.

46. 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. doi: 10.1017/S0950268819001961.

47. O'Hagan JJ, Hernan MA, Walensky RP, Lipsitch M. Apparent declining efficacy in randomized trials: examples of the Thai RV144 HIV vaccine and South African CAPRISA 004 microbicide trials. AIDS. 2012;26:123-6. doi: 10.1097/QAD.0b013e32834e1ce7.

48. Crowcroft NS, Schwartz KL, Savage RD, Chen C, Johnson C, Li Y et al. A call for caution in use of pertussis vaccine effectiveness studies to estimate waning immunity: a Canadian Immunization Research Network Study. Clin Infect Dis. 2020. doi: 10.1093/cid/ciaa518.

49. Halloran ME, Longini IM, Jr., Struchiner CJ. Estimability and interpretation of vaccine efficacy using frailty mixing models. Am J Epidemiol. 1996;144:83-97. doi: 10.1093/oxfordjournals.aje.a008858.

50. Kanaan MN, Farrington CP. Estimation of waning vaccine fficacy. J Am Stat Assoc. 2002;97:389-97. (https://www.jstor.org/stable/3085652?seq=1, accessed 2 March 2021).

51. Zerbo O, Bartlett J, Goddard K, Fireman B, Lewis E, Klein NP. Acellular pertussis vaccine effectiveness over time. Pediatrics. 2019;144. doi: 10.1542/peds.2018-3466.

52. Guidance on developing a national deployment and vaccination plan for COVID-19 vaccines: interim guidance. Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/WHO-2019-nCoV-Vaccine_deployment-2020.1, accessed 2 March 2021).

53. Coronavirus disease (COVID-19) technical guidance: the Unity Studies: early investigation protocols. Geneva: World Health Organization; 2021 (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/early-investigations, accessed 2 March 2021).

54. Evidence used to update the list of underlying medical conditions that increase a person's risk of severe illness from COVID-19. Atlanta (GA): United States Centers for Disease Control and Prevention; 2020 (https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/evidence-table.html, accessed 2 March 2021).

55. Foppa IM, Ferdinands JM, Chaves SS, Haber MJ, Reynolds SB, Flannery B et al. The case test-negative design for studies of the effectiveness of influenza vaccine in inpatient settings. Int J Epidemiol. 2016;45:2052-9. doi: 10.1093/ije/dyw022.

56. 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;[in press]. accessed March 2, 2021.

57. 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. doi: 10.1007/s10654-017-0287-2.

58. Rodrigues L, Kirkwood BR. Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. Int J Epidemiol. 1990;19:205-13. doi: 10.1093/ije/19.1.205.

59. Verani JR, Baqui AH, Broome CV, Cherian T, Cohen C, Farrar JL et al. Case-control vaccine effectiveness studies: preparation, design, and enrollment of cases and controls. Vaccine. 2017;35:3295-302. doi: 10.1016/j.vaccine.2017.04.037.

60. Verani JR, Baqui AH, Broome CV, Cherian T, Cohen C, Farrar JL et al. Case-control vaccine effectiveness studies: data collection, analysis and reporting results. Vaccine. 2017;35:3303-8. doi: 10.1016/j.vaccine.2017.04.035.

61. Olsen SJ, Azziz-Baumgartner E, Budd AP, Brammer L, Sullivan S, Pineda RF et al. Decreased influenza activity during the COVID-19 pandemic - United States, Australia, Chile, and South Africa, 2020. MMWR Morb Mortal Wkly Rep. 2020;69:1305-9. doi: 10.15585/mmwr.mm6937a6.

62. Grippe bulletin hebdomadaire: 03/02/2021. Semaine 04. Santé publique France; 2021 (https://www.santepubliquefrance.fr/content/download/316706/2909158, accessed 2 March 2021).

63. Jackson ML, Rothman KJ. Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness. Vaccine. 2015;33:1313-6. doi: 10.1016/j.vaccine.2015.01.069.

64. 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. doi: 10.1038/s41467-020-19478-2.

65. 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. (http://www.ncbi.nlm.nih.gov/pubmed/3879673, accessed 2 March 2021).

66. Lopez Bernal JA, Andrews N, Amirthalingam G. The use of quasi-experimental designs for vaccine evaluation. Clin Infect Dis. 2019;68:1769-76. doi: 10.1093/cid/ciy906.

67. Basta NE, Halloran ME. Evaluating the effectiveness of vaccines using a regression discontinuity design. Am J Epidemiol. 2019;188:987-90. doi: 10.1093/aje/kwz043.

68. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004;15:615-25. doi: 10.1097/01.ede.0000135174.63482.43.

69. Hernan MA. The hazards of hazard ratios. Epidemiology. 2010;21:13-5. doi: 10.1097/EDE.0b013e3181c1ea43.

70. 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. doi: 10.1093/cid/ciy770.

71. Lipsitch M. Challenges of vaccine effectiveness and waning studies. Clin Infect Dis. 2019;68:1631-3. doi: 10.1093/cid/ciy773.

72. Patel MM, Jackson ML, Ferdinands J. Postlicensure evaluation of COVID-19 vaccines. JAMA. 2020. doi: 10.1001/jama.2020.19328.

73. De Smedt T, Merrall 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. doi: 10.1371/journal.pone.0199180.

74. Lipsitch M, Jha A, Simonsen L. Observational studies and the difficult quest for causality: lessons from vaccine effectiveness and impact studies. Int J Epidemiol. 2016;45:2060-74. doi: 10.1093/ije/dyw124.

75. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 2010;21:383-8. doi: 10.1097/EDE.0b013e3181d61eeb.

76. 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. doi: 10.1093/aje/kwy163.

77. 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. doi: 10.1038/s41591-020-0887-3.

78. Diagnostic testing for SARS-CoV-2: interim guidance. Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2, accessed 2 March 2021).

79. Weiss A, Jellingso M, Sommer MOA. Spatial and temporal dynamics of SARS-CoV-2 in COVID-19 patients: a systematic review and meta-analysis. EBioMedicine. 2020;58:102916. doi: 10.1016/j.ebiom.2020.102916.

80. FIND evaluation update: SARS-CoV-2 molecular diagnostics. FIND; 2021 (https://www.finddx.org/covid-19-old/sarscov2-eval-molecular/, accessed 2 March 2021).

81. Hodgson SH, Mansatta K, Mallett G, Harris V, Emary KRW, Pollard AJ. What defines an efficacious COVID-19 vaccine? A review of the challenges assessing the clinical efficacy of vaccines against SARS-CoV-2. Lancet Infect Dis. 2020. doi: 10.1016/s1473-3099(20)30773-8.

82. SARS-CoV-2 genomic sequencing for public health goals: interim guidance, 8 January 2021. Geneva: World Health Organization; 2021 (https://www.who.int/publications/i/item/WHO-2019-nCoV-genomic_sequencing-2021.1, accessed 2 March 2021).

83. Skowronski DM, Chambers C, Sabaiduc S, De Serres G, Winter AL, Dickinson JA et al. Integrated sentinel surveillance linking genetic, antigenic, and epidemiologic monitoring of influenza vaccine-virus relatedness and effectiveness during the 2013-2014 influenza season. J Infect Dis. 2015;212:726-39. doi: 10.1093/infdis/jiv177.

84. WHO reference laboratories providing confirmatory testing for COVID-19. Geneva: World Health Organization; 2020 (https://www.who.int/publications/m/item/who-reference-laboratories-providing-confirmatory-testing-for-covid-19, accessed 2 March 2021).

85. O'Neill RT. On sample sizes to estimate the protective efficacy of a vaccine. Stat Med. 1988;7:1279-88. doi: 10.1002/sim.4780071208.

86. Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ. Modern epidemiology. 4th ed. Philadelpha (PA): Lippincott Williams & Wilkins; 2021.

87. Mansournia MA, Hernan MA, Greenland S. Matched designs and causal diagrams. Int J Epidemiol. 2013;42:860-9. doi: 10.1093/ije/dyt083.

88. Mickey RM, Greenland S. The impact of confounder selection criteria on effect estimation. Am J Epidemiol. 1989;129:125-37. doi: 10.1093/oxfordjournals.aje.a115101.

89. Greenland S. Modeling and variable selection in epidemiologic analysis. Am J Public Health. 1989;79:340-9. doi: 10.2105/ajph.79.3.340.

90. Johnson & Johnson announces single-shot Janssen COVID-19 vaccine candidate met primary endpoints in interim analysis of its Phase 3 ENSEMBLE Trial. Johnson & Johnson; 2021 (https://www.jnj.com/johnson-johnson-announces-single-shot-janssen-covid-19-vaccine-candidate-met-primary-endpoints-in-interim-analysis-of-its-phase-3-ensemble-trial#:~:text=Among%20all%20participants%20from%20different,as%20early%20as%20day%2014, accessed 2 March 2021).

91. Munoz FM, Cramer JP, Dekker CL, Dudley MZ, Graham BS, Gurwith M et al. Vaccine-associated enhanced disease: case definition and guidelines for data collection, analysis, and presentation of immunization safety data. Brighton Collaboration; 2020 (https://brightoncollaboration.us/wp-content/uploads/2020/10/JVAC-S-20-02520_mzd.pdf, accessed 2 March 2021).

92. Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. Int J Epidemiol. 1999;28:1-9. doi: 10.1093/ije/28.1.1.

93. Friedenreich CM. Methods for pooled analyses of epidemiologic studies. Epidemiology. 1993;4:295-302. doi: 10.1097/00001648-199307000-00004.

94. Managing heterogeneity when pooling data from different surveillance systems. Stockholm: European Centre for Disease Prevention and Control; 2019 (https://www.ecdc.europa.eu/en/publications-data/managing-heterogeneity-when-pooling-data-different-surveillance-systems, accessed 2 March 2021).

95. Valenciano M, Kissling E, Larrauri A, Nunes B, Pitigoi D, O'Donnell J et al. Exploring the effect of previous inactivated influenza vaccination on seasonal influenza vaccine effectiveness against medically attended influenza: results of the European I-MOVE multicentre test-negative case-control study, 2011/2012-2016/2017. Influenza Other Respir Viruses. 2018;12:567-81. doi: 10.1111/irv.12562.

96. El Omeiri N, Azziz-Baumgartner E, Thompson MG, participants RE-in, Clara W, Cerpa M et al. Seasonal influenza vaccine effectiveness against laboratory-confirmed influenza hospitalizations - Latin America, 2013. Vaccine. 2018;36:3555-66. doi: 10.1016/j.vaccine.2017.06.036.

97. I-Move: Influenza - Monitoring vaccine effectiveness in Europe. I-MOVE; 2021 (https://www.imoveflu.org/, accessed 2 March 2021).

98. Global Influenza Surveillance and Response System (GISRS). Geneva: World Health Organization; 2021 (https://www.who.int/influenza/gisrs_laboratory/en/, accessed 2 March 2021).

99. Global Influenza Hospital Surveillance Network. 2021 (https://www.gihsn.org/, accessed 2 March 2021).

100. Red para la Evaluación de la Efectividad de la Vacuna en Latino América y el Caribe – influenza (REVELAC-i). Pan American Health Organization; 2021 (https://www.paho.org/revelac-i/, accessed 2 March 2021).

101. Protocol for assessment of potential risk factors for 2019-novel coronavirus (COVID-19) infection among health care workers in a health care setting. Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/protocol-for-assessment-of-potential-risk-factors-for-2019-novel-coronavirus-(2019-ncov)-infection-among-health-care-workers-in-a-health-care-setting, accessed 2 March 2021).

102. COVID-19 vaccines: safety surveillance manual. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/bitstream/handle/10665/338400/9789240018280-eng.pdf?sequence=1&isAllowed=y, accessed 2 March 2021).

103. Research Ethics Review Committee (ERC). Geneva: World Health Organization; 2021 (https://www.who.int/groups/research-ethics-review-committee/guidelines-on-submitting-research-proposals-for-ethics-review/templates-for-informed-consent-forms, accessed 2 March 2021).

104. 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. doi: 10.1016/S1473-3099(16)00129-8.

105. 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. doi: 10.1586/14760584.2014.966695.

106. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche 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. doi: 10.1016/S0140-6736(07)61602-X.

107. Immunization and Vaccine-related Implementation Research Advisory Committee (IVIR-AC): summary of conclusions and recommendations, 1-2 February 2017 meeting. Wkly Epidemiol Rec. 2017;92:181-8. (https://apps.who.int/iris/bitstream/handle/10665/255008/WER9215.pdf;jsessionid=DF9852BC11FEC876A5CDD1C26104B9A1?sequence=1, accessed 2 March 2021).

108. Enhancing the QUAlity and Transparency Of health Research. Equator Network; 2021 (https://www.equator-network.org/reporting-guidelines/strobe/, accessed 2 March 2021).

109. COVID-19 data collection. Harvard Dataverse; 2021 (https://dataverse.harvard.edu/, accessed 2 March 2021).

110. COVID-19 Therapeutic Trial Synopsis. Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/covid-19-therapeutic-trial-synopsis, accessed 2 March 2021).

111. Development and licensure of vaccines to prevent COVID-19: guidance for industry. U.S. Food & Drug Administration; 2020 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/development-and-licensure-vaccines-prevent-covid-19, accessed 2 March 2021).

Annex 1: Examples of vaccine effectiveness protocols

These are some protocols and additional guidance documents that might be of use for countries in developing their COVID-19 VE protocols.

Protocol/guidance	Study description	Link
Influenza - Monitoring Vaccine Effectiveness in Europe (I-MOVE): COVID-19 vaccine effectiveness at primary care level in Europe (generic protocol)	Test-negative design study among persons with symptomatic COVID-19 seeking primary care	https://www.imoveflu.org/wp-content/uploads/2021/03/I-MOVE-COVID-19-primary-care-COVID-19-vaccine-effectiveness-protocol-v2.2.pdf ()
Influenza - Monitoring Vaccine Effectiveness in Europe (I-MOVE): European study of COVID-19 vaccine effectiveness against hospitalized SARI patients laboratory confirmed with SARS-CoV-2 (draft generic protocol)	Test-negative design study among hospitalized SARI patients	https://www.imoveflu.org/wp-content/uploads/2021/03/08feb2021_draft_generic_VE_protocol_hospital-based_COVID-19_v07.pdf ()
Cohort study to measure COVID-19 vaccine effectiveness among health workers in the WHO European Region: guidance document	Cohort study among health care workers	https://apps.who.int/iris/handle/10665/340217 ()
WHO-Europe Guidance Document: COVID-19 vaccine effectiveness against severe acute respiratory infections (SARI) hospitalizations associated with laboratory-confirmed SARS-CoV-2	Test-negative design study among hospitalized SARI patients	Link pending
Unity Study: Vaccine Effectiveness	Test negative design study	Coming Soon
Unity Study: Household transmission investigation protocol for 2019-novel coronavirus (COVID-19) infection	Household transmission study Note that changes in line with this guidance document need to be made to the protocol (e.g. collection of vaccination details)	https://www.who.int/publications-detail-redirect/household-transmission-investigation-protocol-for-2019-novel-coronavirus-(2019-ncov)-infection ()
Unity Study: A prospective cohort study investigating maternal, pregnancy and neonatal outcomes for women and neonates infected with SARS-CoV-2	Cohort study of pregnant women whereby one could calculate VE in this population Note that changes in line with this guidance document need to be made to the protocol (e.g. collection of vaccination details)	https://www.who.int/publications/m/item/a-prospective-cohort-study-investigating-maternal-pregnancy-and-neonatal-outcomes-for-women-and-neonates-infected-with-sars-cov-2 ()

Annex 2: Sample size

Minimum number of cases and controls to detect for a specified vaccine effectiveness, estimated vaccination coverage in the population under evaluation, with 1–4 controls per case, with a precision of: ± 10%; and ± 5%, and a type 1 error rate of 0.05.

Table A2.1 Precision of estimate of ± 10%

Vaccine effectiveness	Vaccination coverage in the population being studied	1:1 cases to controls		1:2 cases to controls		1:3 cases to controls		1:4 cases to controls	
		No. cases	No. controls	No. cases	No. controls	No. cases	No. controls	No. cases	No. controls
50%	20%	1594	1594	1290	2580	1188	3564	1138	4552
	30%	1133	1133	902	1804	825	2475	786	3144
	40%	925	925	722	1444	655	1965	621	2484
	50%	828	828	633	1266	568	1704	536	2144
	60%	803	803	601	1202	533	1599	499	1996
	70%	855	855	624	1248	546	1638	508	2032
	80%	1047	1047	742	1484	641	1923	590	2360
	90%	1736	1736	1195	2390	1015	3045	925	3700
60%	20%	1152	1152	956	1912	890	2670	858	3432
	30%	801	801	652	1304	602	1806	577	2308
	40%	639	639	509	1018	465	1395	443	1772
	50%	559	559	433	866	392	1176	371	1484
	60%	530	530	399	798	355	1065	333	1332
	70%	550	550	401	802	351	1053	326	1304
	80%	658	658	462	924	396	1188	364	1456
	90%	1066	1066	718	1436	602	1806	543	2172
70%	20%	776	776	664	1328	627	1881	609	2436
	30%	526	526	441	882	412	1236	398	1592
	40%	408	408	333	666	308	924	296	1184
	50%	346	346	274	548	250	750	238	952
	60%	317	317	243	486	218	654	205	820
	70%	319	319	234	468	205	615	191	764
	80%	369	369	257	514	220	660	201	804
	90%	580	580	381	762	315	945	282	1128
80%	20%	470	470	418	836	401	1203	392	1568
	30%	308	308	268	536	255	765	248	992
	40%	229	229	195	390	183	549	178	712
	50%	186	186	153	306	142	426	137	548
	60%	163	163	129	258	117	351	111	444
	70%	156	156	116	232	103	309	97	388
	80%	171	171	120	240	102	306	94	376
	90%	257	257	165	330	134	402	119	476
90%	20%	239	239	224	448	219	657	216	864
	30%	150	150	138	276	134	402	132	528
	40%	106	106	95	190	92	276	90	360
	50%	80	80	70	140	67	201	65	260
	60%	65	65	54	108	51	153	49	196
	70%	56	56	45	90	41	123	39	156
	80%	56	56	40	80	35	105	32	128
	90%	75	75	48	96	39	117	34	136

Table A2.2 Precision of estimate of ± 5%

Vaccine effectiveness	Vaccination coverage in the population being studied	1:1 cases to controls		1:2 cases to controls		1:3 cases to controls		1:4 cases to controls	
		No. cases	No. controls	No. cases	No. controls	No. cases	No. controls	No. cases	No. controls
50%	20%	6312	6312	5107	10214	4706	14118	4505	18020
	30%	4488	4488	3570	7140	3264	9792	3111	12444
	40%	3662	3662	2859	5718	2591	7773	2458	9832
	50%	3277	3277	2506	5012	2249	6747	2120	8480
	60%	3180	3180	2377	4754	2110	6330	1976	7904
	70%	3387	3387	2469	4938	2163	6489	2010	8040
	80%	4144	4144	2939	5878	2538	7614	2337	9348
	90%	6874	6874	4733	9466	4019	12057	3662	14648
60%	20%	4535	4535	3763	7526	3506	10518	3377	13508
	30%	3156	3156	2567	5134	2371	7113	2273	9092
	40%	2517	2517	2002	4004	1831	5493	1745	6980
	50%	2200	2200	1706	3412	1541	4623	1459	5836
	60%	2085	2085	1570	3140	1398	4194	1312	5248
	70%	2167	2167	1579	3158	1382	4146	1284	5136
	80%	2589	2589	1817	3634	1559	4677	1431	5724
	90%	4199	4199	2826	5652	2368	7104	2140	8560
70%	20%	3023	3023	2587	5174	2442	7326	2369	9476
	30%	2048	2048	1715	3430	1605	4815	1549	6196
	40%	1587	1587	1296	2592	1199	3597	1151	4604
	50%	1345	1345	1066	2132	973	2919	926	3704
	60%	1234	1234	943	1886	846	2538	798	3192
	70%	1241	1241	909	1818	798	2394	743	2972
	80%	1436	1436	1000	2000	854	2562	781	3124
	90%	2259	2259	1484	2968	1225	3675	1096	4384
80%	20%	1776	1776	1580	3160	1514	4542	1482	5928
	30%	1162	1162	1013	2026	963	2889	938	3752
	40%	866	866	735	1470	692	2076	670	2680
	50%	703	703	578	1156	536	1608	515	2060
	60%	615	615	485	970	441	1323	419	1676
	70%	588	588	439	878	389	1167	364	1456
	80%	647	647	451	902	385	1155	353	1412
	90%	971	971	622	1244	506	1518	448	1792
90%	20%	801	801	749	1498	732	2196	724	2896
	30%	500	500	461	922	448	1344	441	1764
	40%	353	353	318	636	307	921	301	1204
	50%	268	268	234	468	223	669	218	872
	60%	216	216	181	362	170	510	164	656
	70%	188	188	148	296	135	405	128	512
	80%	185	185	134	268	116	348	108	432
	90%	251	251	159	318	128	384	113	452

Annex 3: Possible case definitions; inclusion and exclusion criteria

A3.1 Possible case definitions

For TND and traditional case-control studies, having a strict case definition to determine enrolment eligibility is extremely important as this helps to decrease some biases that can arise from variability in clinical diagnosis. An enrolment case definition should be clear and be simple to apply by sites. Enrolment case definitions for VE evaluations, however, should not be used to guide clinical management. A variety of possible definitions are provided below, and each evaluation should use or modify the definition that fits the objective of the evaluation and can be applied in the country setting.

Symptomatic COVID-19 disease

	Source of definition	Definition
Suspected COVID-19 case definition	Modification of WHO COVID-19 surveillance guidelines *(40)*	A person who has had the following symptoms within the last 10 days: • acute onset of fever AND cough; OR • acute onset of ANY THREE OR MORE of the following signs or symptoms: fever, cough, general weakness/fatigue, headache, myalgia, sore throat, coryza, dyspnoea, anorexia/nausea/vomiting, diarrhoea, altered mental status.
Influenza-like illness (ILI) case definition	WHO surveillance case definition for ILI *(39)*	A person with an acute respiratory infection with measured fever of ≥ 38°C, cough, with onset within the last 10 days.

Severe COVID-19 disease

If the objective of the evaluation is to assess VE against severe disease, severity must be defined within the confines of what is feasible in the evaluation setting. Different severity scales exist and are provided below.

WHO SEVERITY SCALE

As defined by WHO's *Clinical management of COVID-19: interim guidance (38).* Note that this severity scale does not include death as these come from clinical treatment guidelines; death can be considered the most severe form of disease.

Mild disease:

Someone who is symptomatic meeting the case definition for COVID-19 without evidence of viral pneumonia or hypoxia.

Moderate disease:

- Adolescent or adult with clinical signs of pneumonia (fever, cough, dyspnoea, fast breathing) but no signs of severe pneumonia, including SpO_2 ≥ 90% on room air.
- Child with clinical signs of non-severe pneumonia (cough or difficulty breathing + fast breathing and/or chest indrawing) and no signs of severe pneumonia.
  - fast breathing (in breaths/min): < 2 months: ≥ 60; 2–11 months: ≥ 50; 1–5 years: ≥ 40.

Severe disease:

- Adolescent or adult with clinical signs of pneumonia (fever, cough, dyspnoea, fast breathing) plus one of the following: respiratory rate > 30 breaths/min; severe respiratory distress; or SpO_2 < 90% on room air.
- Child with clinical signs of pneumonia (cough or difficulty in breathing) + at least one of the following:
  - central cyanosis or SpO_2 < 90%; severe respiratory distress (e.g. fast breathing, grunting, very severe chest indrawing); general danger sign: inability to breastfeed or drink, lethargy or unconsciousness, or convulsions;
  - fast breathing (in breaths/min): < 2 months: ≥ 60; 2–11 months: ≥ 50; 1–5 years: ≥ 40.

While the diagnosis can be made on clinical grounds; chest imaging (radiograph, CT scan, ultrasound) may assist in diagnosis and identify or exclude pulmonary complications.

Critical disease:

A person with ARDS, sepsis or septic shock. Full details are available in WHO's *Clinical management of COVID-19: interim guidance (38)*.

It is recommended to include death if using this definition as the most severe form of disease.

WHO'S SEVERITY CRITERIA FROM COVID-19 THERAPEUTIC TRIAL SYNOPSIS (SOLIDARITY TRIAL) *(110)*

Patient status	Description of severity	Score
Ambulatory	No limitation of activities	1
	Limitation of activities	2
Hospitalized mild disease	Hospitalize, no oxygen therapy	3
	Oxygen by mask or nasal cannula	4
Hospitalized severe disease	Non-invasive ventilation or high-flow oxygen	5
	Intubation and mechanical ventilation	6
	Ventilation and additional organ support (e.g. pressors, extracorporeal membrane oxygenation [ECMO])	7
Dead	Death	8

UNITED STATES FOOD & DRUG ADMINISTRATION'S SEVERITY DEFINITION

This is suggested for use in Phase III clinical trials of COVID-19 vaccines *(111)*.

Severe COVID-19 is defined as virologically confirmed SARS-CoV-2 infection with any of the following:

- clinical signs at rest indicative of severe systemic illness (respiratory rate ≥ 30 per minute, heart rate ≥ 125 per minute, SpO_2 ≤ 93% on room air at sea level or PaO_2/FiO_2 < 300 mm Hg);
- respiratory failure (defined as needing high-flow oxygen, non-invasive ventilation, mechanical ventilation or ECMO);
- evidence of shock (SBP < 90 mm Hg, DBP < 60 mm Hg, or requiring vasopressors);
- significant acute renal, hepatic, or neurologic dysfunction;
- admission to an ICU;
- death.

A3.2 Suggested inclusion criteria for cases and controls

Suggested inclusion criteria for cases and controls using several study designs are provided below.

For cases (all designs) and TND controls

- eligible to have received COVID-19 vaccine;
- admitted or presenting to health facility (outpatient/inpatient/emergency room) for acute medical illness and meets enrolment case definition;
- underwent testing for SARS-CoV-2 with a test with ≥ 85% sensitivity and ≥ 98% specificity for the currently circulating variants;
- if applicable, able to provide informed consent (or has a proxy able to provide consent);
- specimen collected within 10 days of symptom onset.

Health facility controls

- eligible to have received COVID-19 vaccine;
- admitted or presenting to health facility for non-COVID-19 like illness:
  - examples include persons hospitalized due to non-respiratory illnesses, such as urinary tract infection, skin/soft tissue infections, trauma, surgery, obstetrics;
- no positive SARS-CoV-2 test during the current illness;
- if applicable, able to provide informed consent (or has a proxy able to provide consent).

A3.3 Potential exclusion criteria

Potential exclusion criteria are provided below. Some of these criteria help to decrease some biases but can also introduce other biases if too many persons are excluded who meet the suspected case definition. One can choose to exclude some persons a priori; in other instances, one may choose to conduct a sensitivity analysis using various criteria to evaluate their influence on the results.

- Those not consenting, who decline or are unable to be interviewed by survey staff or those without caregivers (if informed consent required by local requirements).

- Not in target group for COVID-19 vaccine (e.g. outside age range).

- If patient has a contraindication to specimen collection, specimens cannot be collected, or test results are unavailable.

- Patients that may not be representative of the source population (e.g. patients who were transferred from an outside facility).

- Were vaccinated with their first dose within 14 days of symptom onset (exclusion from primary analysis but could include in secondary analyses).

- Were vaccinated with their second dose within 7–14 days of symptom onset (exclusion from primary analysis but could include in secondary analyses).

Annex 4: Reporting elements

STROBE checklist *(106)* and recommended additional elements for reporting COVID-19 VE studies

Section/topic	STROBE Item no.	STROBE	COVID-19 VE studies
Title and abstract			
Title/abstract	1	Indicate the study's design with a commonly used term in the title or the abstract	• Specify study design (e.g. case-control, TND or cohort)
		Provide in the abstract an informative and balanced summary of what was done and what was found	• Report vaccine type(s), outcome, target vaccine groups evaluated, study location, VE and 95% CI
Introduction			
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	• 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	• 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	• 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	• Describe the enrolment 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 • Report time period when data were collected
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	• Report specific clinical case definition used for enrolment • Report definition of severity used • Describe eligible study population in terms of age and vaccine target groups (e.g. health workers, chronic medical conditions) and exclusion criteria

STROBE checklist *(106)* **and recommended additional elements for reporting COVID-19 VE studies** continued

Section/topic	STROBE Item no.	STROBE	COVID-19 VE studies
Variables	**7**	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers Give diagnostic criteria, if applicable	*COVID-19 vaccine variables* • 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 *COVID-19 outcomes* • 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 enrolment • Explain how possible vaccine reactions were handled in TND studies (e.g. exclude recent vaccine recipients tested for possible febrile reaction to vaccine) *Covariates* • 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 of 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	*COVID-19 vaccine* • 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) *COVID-19 outcomes* • 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. flocked), 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 enrolment 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	• 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	• Adjust sample size calculation to expected COVID-19 incidence and estimated VE from clinical trial
Quantitative variables	**11**	Explain how quantitative variables were handled in the analyses If applicable, describe which groups were chosen and why	• 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

STROBE checklist *(106)* and recommended additional elements for reporting COVID-19 VE studies continued

Section/topic	STROBE Item no.	STROBE	COVID-19 VE studies
Statistical methods	**12**	Describe all statistical methods, including those used to control for confounding	• Describe the specific regression method used (e.g. logistic regression) and confidence limits methodology • Report the time periods for which data were analysed 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	**11**	Describe any methods used to examine subgroups and interactions	• Describe any analyses of subgroups (e.g. age groups, chronic conditions, health workers) • Describe interactions assessed (e.g. prior COVID-19 infection)
		Explain how missing data were addressed	• 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)
		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	• In case-control studies, if more than one control group enrolled, explain rationale
		Describe any sensitivity analyses	• For example, excluding verbal reports of vaccination; limited to positive test within 72 hours of enrolment; limited to PCR+ only (if rapid antigen tests included)
Other			• Indicate if and where study protocol and/or study data are publicly available
Results			
Participants	**13**	a) 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 analysed) b) Give reasons for non-participation at each stage c) Consider use of a flow diagram	
Descriptive data	**14**	a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders	• 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
		b) Indicate number of participants with missing data for each variable of interest	
		c) Cohort study: summarize follow-up time (e.g. average and total amount)	
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	• Describe number/percentage of tests which were PCR, rapid antigen test, other • Report COVID-19 genomic information among vaccine failures, if available; particularly variants of concern

STROBE checklist *(106)* and recommended additional elements for reporting COVID-19 VE studies continued

Section/topic	STROBE Item no.	STROBE	COVID-19 VE studies
Main results	**16**	a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% CI); make clear which confounders were adjusted for and why they were included	• 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
		b) Report category boundaries when continuous variables were categorized	
		c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	**17**	Report other analyses done, e.g. analyses of subgroups and interactions, and sensitivity analyses	• 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 of 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	• 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	• 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	• Was baseline seroprevalence different from other settings? • Predominant viral variant found in other settings?
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	

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

World Health Organization
Avenue Appia 20
1211 Geneva 27
Switzerland
www.who.int