Guidance on conducting vaccine effectiveness evaluations in the setting of new SARS-CoV-2 variants

INTERIM GUIDANCE
22 JULY 2021

Addendum to Evaluation of COVID-19 vaccine effectiveness: Interim guidance
Guidance on conducting vaccine effectiveness evaluations in the setting of new SARS-CoV-2 variants

INTERIM GUIDANCE
22 JULY 2021

Addendum to Evaluation of COVID-19 vaccine effectiveness: Interim guidance
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 document will expire 2 years after the date of publication.

© World Health Organization 2021. Some rights reserved. This work is available under the CC BY-NC-SA 3.0 IGO licence.

WHO/2019-nCoV/vaccine_effectiveness/variants/2021.1

Design and layout by L'IV Com Sàrl, Switzerland
## Contents

Acknowledgements ................................................................. iv  
Abbreviations ........................................................................ v  
1. Introduction ........................................................................... 1  
2. Surveillance for new variants ................................................... 2  
3. Early triggers of reduced vaccine effectiveness against new variants ........................................... 3  
   3.1 Vaccine breakthrough cases and new variants ................................................................. 3  
   3.2 Changing epidemiology of COVID-19 as possible indication of reduced vaccine effectiveness against new variants ...................................................... 4  
   3.3 Screening method assessment using breakthrough cases to assess if vaccine effectiveness might be lower than expected in the setting of new variants ................. 4  
4. Case-only approaches to assessing vaccine effectiveness from new variants ............................ 9  
   4.1 Comparison of variant prevalence in vaccinated and unvaccinated cases ................... 9  
   4.2 Sieve analysis ....................................................................... 10  
5. Full vaccine effectiveness evaluations for new variants .......................................................... 12  
   5.1 Evaluation of COVID-19 VE when new variant is co-circulating ............................... 12  
     A. When the majority of cases have genomic characterization ........................................... 12  
     B. When a majority of cases do not have genomic characterization ................................. 13  
   5.2 Evaluation of COVID-19 VE when new variant is the predominant circulating strain ........ 14  
6. Analytic considerations for genomic sequencing when estimating vaccine effectiveness for new variants ................................................................. 15  
7. Bias in vaccine effectiveness studies of new variants ................................................................. 16  
8. Conclusion .............................................................................. 17  
References .................................................................................. 18
Acknowledgements

This addendum was written by the following people, listed alphabetically:

Natasha Crowcroft (Department of Immunization, Vaccines, and Biologicals, WHO)
Daniel Feikin (Department of Immunization, Vaccines, and Biologicals, WHO)
Brendan Flannery (US CDC)
Rebecca Kahn (Harvard University)
Esther Kissling (Epiconcept)
Maria Knoll (Johns Hopkins Bloomberg School of Public Health)
Marc Lipsitch (Harvard University)
Olivier Le Polain de Waroux (Health Emergencies Programme, WHO)
Walt Orenstein (Emory University)
Minal Patel (Department of Immunization, Vaccines, and Biologicals, WHO)
Morgane Rolland (Walter Reed Army Institute of Research)
Marta Valenciano (Epiconcept)
Jennifer Verani (US CDC)

We thank the following people at WHO for review and comment of an earlier draft: Brett Archer, Boris Pavlin, Lorenzo Subissi, Jillian Sacks and Mick Mulders.
## Abbreviations

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

The World Health Organization (WHO) has released working case definitions for SARS-CoV-2 variants of interest (VOIs) (1) and variants of concern (VOCs) (1). VOCs are SARS-CoV-2 variants with an increase in transmissibility or detrimental change in COVID-19 epidemiology, an increase in virulence or change in clinical disease presentation, or a decrease in effectiveness of public health and social measures or available diagnostics, vaccines or therapeutics. As the pandemic progresses new variants will likely emerge, especially in areas and groups with high incidence and low vaccine coverage. Evidence that vaccines may be less protective against a specific variant may be suggested by genomic and structural biology studies, animal studies and in-vitro neutralization testing. Lower effectiveness of a vaccine in protecting against infection and disease from a variant in humans, however, provides the strongest evidence. Rarely, a new variant may be circulating at the time of a randomized controlled trial of a new vaccine, as was the case for the Beta (B1.351) variant in South Africa for several vaccines (2–4). Lower vaccine efficacy against a variant than against other strains within a randomized trial provides the highest quality evidence. More often, however, new variants will emerge after clinical trials are conducted or in different locations, and most epidemiological data on vaccine performance against new variants will come from observational studies of vaccine effectiveness (VE).

2. Surveillance for new variants

Systematic epidemiological surveillance must be coupled with genomic characterization of viruses to evaluate VE against new variants. WHO has published interim guidance on surveillance for COVID-19 and genomic sequencing of virus isolates (6, 7). WHO is issuing updated guidance on surveillance for SARS-CoV-2 variants, which emphasizes the addition of genomic sequencing to routine surveillance, sampling of cases for genomic sequencing, reporting of new variants and implications of circulating variants on public health and transmission control measures (8). Routine ongoing surveillance for COVID-19 can produce signals indicating the possible presence a new VOC/VOI circulating in the population. This may include increased COVID-19 incidence, hospitalization or mortality rates among vaccinated cohorts or increased frequency and size of COVID-19 outbreaks, particularly in highly vaccinated populations.

Additionally, increasing rates of invalid test results or test failure could be an indication that the virus includes mutations that render it unable to be detected by diagnostic tests. Similarly, unexpected test results, for example negative antigen test in a person with a high viral load by nucleic acid testing could indicate that a variant is no longer detectable by the antigen-detection assay. Though there are many factors that influence test performance, one aspect can be the presence of variants.

WHO has published two interim guidance documents for SARS-CoV-2 genomic sequencing to assist in identifying mutations among circulating viruses (6, 9). Whole genome sequencing or targeted gene sequencing particularly of the S gene, are useful to identify and characterize virus evolution and the circulation of variants. As well, nucleic acid amplification tests (NAAT), such as PCR, exist which can specifically detect single nucleotide polymorphisms (SNPs)/mutations that are characteristic of VOI/VOC. Depending on the context, positive results from such mutation screening assays may be indicative of a specific variant but are not confirmatory. Therefore, all or at least a subset of positive samples should be referred for sequencing to confirm the presence of a specific variant. Various sampling strategies for genomic characterization of samples from routine surveillance have been proposed, but in all instances the sampling approach should be systematic, representative for age groups, geography and clinical severity, continuous, and timely so that variants can be identified early and in proportion to their distribution in the population. Collection of metadata on samples to be sequenced from cases is important and should include date and location, patient characteristics, whether a case is imported or locally acquired, clinical presentation and, if available, measures of severity of the case. Case-based surveillance is also needed for determining case vaccination status. For breakthrough cases, vaccine type and date of vaccination should be included to evaluate whether immune escape from vaccine immunity might be occurring.

The four VOCs replaced the other circulating variants within a few months in countries where they were first detected, and this pattern is repeating in other countries. A recent analysis of these variants shows that they have a 29–97% increase in transmissibility compared to the original strains (10). However, whether introduction of these VOCs into additional countries will necessarily lead to rapid replacement of circulating viruses is not clear and likely depends on the prevalence of other variants. For example, Beta circulated at low levels in Israel and then disappeared in the context of Alpha predominance and high vaccine coverage; while in Qatar, Beta has become the dominant strain and, in the United Kingdom, Delta has replaced Alpha (11, 12).
3. Early triggers of reduced vaccine effectiveness against new variants

3.1 Vaccine breakthrough cases and new variants

Vaccine breakthrough cases are expected to occur regardless of virus strain because no vaccine is 100% effective. Breakthrough cases should not necessarily be seen as a failure of the vaccine. However, vaccine breakthrough cases may signal reduced vaccine effectiveness against emergent viruses. An asymptomatic breakthrough infection is defined as the detection of SARS-CoV-2 RNA or antigen in a respiratory specimen collected from a person without COVID-like symptoms ≥14 days after they have completed all recommended doses of the vaccine series. A symptomatic breakthrough case is defined as the detection of SARS-CoV-2 RNA or antigen in a respiratory specimen collected from a person with COVID-like symptoms ≥14 days after they have completed all recommended doses of the vaccine series. Patterns of vaccine breakthrough cases caused by new variants may depend on individual immune response to vaccination as discussed below.

A. Primary vaccine failure [insufficient immune response for protection]. For unknown reasons, some vaccinations may entirely fail to elicit a protective immune response in a proportion of vaccinated individuals. This phenomenon is most often associated with live virus vaccines, like measles, but can occur with any vaccine whether live or killed. Primary vaccine failure can occur for other reasons; for example due to breaks in the cold chain, damage to the vaccine before administration, inappropriate administration or underlying immunodeficiency in the recipient. For many persons with primary vaccine failure, the cause is unknown. This mode of failure would apply regardless of what variant an individual is exposed to, because it depends on a failure to generate an immune response.

B. Vaccine modified disease. Individual immune responses may not prevent infection or disease but may result in fewer symptoms or less severe outcomes. Individual immune responses may initially be strong enough to prevent any infection or disease but may wane over time and prevent only severe disease that occurs sometime after vaccination. However, a new variant to which the vaccine’s initial immune response is lower might escape immune response to a vaccine sooner after vaccination, resulting in increased severity of disease among breakthrough infections occurring sooner after vaccination.

C. “Exposure threshold” vaccines. The effectiveness of vaccines and frequency of vaccine breakthrough cases may vary according to the level of exposure. Hence the investigation of vaccine breakthrough cases should include the context, closeness of contact, local epidemiology and other environmental factors that may influence the exposure level of individuals or groups. New variants of concern appear to be more transmissible, in part due to an increased viral load and replication (13, 14). Moreover, the titres of neutralization antibodies needed to protect against some new variants (e.g. Beta and Delta) are higher (15). These factors might allow new variants to cause breakthrough infection and/or disease at lower exposure levels; although this would be hard to measure.

D. Secondary vaccine failure. Vaccination may provide protection for some time before vaccines no longer protect against disease due to waning immunity. While decay of antibody concentrations and cell-mediated immunity likely occurs over months to years, the risk of being infected may be threshold related. However, the risk of becoming infected may be threshold related. Antibody concentrations above a threshold that correlate with protection may prevent infection or disease, while concentrations...
below the threshold may not provide protection. From an epidemiological perspective, vaccine breakthroughs would start to increase at a certain time after vaccination of a population, which may occur earlier among age groups first vaccinated in campaigns or with targeted strategies (e.g., elderly persons, health-care workers). As antibody neutralization might be lower against new variants, loss of protection and vaccine breakthrough cases due to emerging variants may be observed sooner than with vaccine-reference strains although this has yet to be documented. Of note, since protection relies on both humoral and cellular immunity, lower antibody levels do not necessarily rule out immunity as a whole since cellular immunity might still be providing protection, particularly against severe disease.

3.2 Changing epidemiology of COVID-19 as possible indication of reduced vaccine effectiveness against new variants

Besides an increase in vaccine breakthrough cases, changes in COVID-19 epidemiology might provide an early signal of circulation of a new variant virus associated with decreased vaccine effectiveness. Within weeks of reaching high vaccination coverage in a population, substantial decreases in disease incidence may be observed in target groups, as was observed in the rapid rollout of vaccine in Israel (16). Unexpected slowing or reversals of downward trends may signal reduced VE against circulating viruses, particularly in the context of emerging variants. An increase in COVID-19 hospitalizations or deaths in a highly vaccinated population could also signal reduced VE, as could disease outbreaks in highly vaccinated congregate living communities, such as long-term care facilities or military barracks. In all of these situations, genomic sequencing should be part of the investigation to determine if a new variant has become more prevalent in the population.

3.3 Screening method assessment using breakthrough cases to assess if vaccine effectiveness might be lower than expected in the setting of new variants

Evaluating breakthrough cases alone cannot quantify VE – information about vaccination coverage in the population is needed. The screening method allows estimation of VE by comparing the vaccination coverage among cases to that in the source population in which the cases were identified.

The screening method serves as a rapid tool to assess whether a vaccine is performing as expected, including against new variants. Lower VE estimates than expected based on clinical trials or observational studies should trigger a more rigorous VE evaluation. Although lower than expected crude VE could be due to a new variant, other potential causes should also be investigated, as outlined in the interim VE guidance (see Box 1).

Only two data points are needed to calculate a crude VE using the screening method: the percentage of total cases occurring in fully vaccinated persons (i.e., breakthrough cases as a proportion of all cases); and the vaccination coverage in the population. The screening method is therefore relatively easy to perform and inexpensive (17). Figure 1 shows the relationship between the percentage of the cases who are vaccinated, the percentage of the source population that is vaccinated and the crude VE. If data on cases caused by variants are available, a variant-specific crude VE can be calculated using the screening method by substituting the percentage of all cases who are vaccinated with the percentage of variant cases who are vaccinated.
BOX 1. POTENTIAL REASONS THAT CRUDE VE ESTIMATES DIFFER FROM EXPECTED

<table>
<thead>
<tr>
<th>VE estimate valid</th>
<th>VE estimate not valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population being studied has different VE for epidemiologic or biological reasons (e.g. VOC, age, underlying condition)</td>
<td>Inaccuracies in the data being used (e.g. ascertainment of vaccination status of cases differs from estimated coverage in population)</td>
</tr>
<tr>
<td>Vaccine mishandling</td>
<td>Biases (e.g. selection bias due to missing data on cases, prioritization for testing/sequencing among vaccinees or surveillance for vaccine failures)</td>
</tr>
<tr>
<td>Systematic error in vaccine administration</td>
<td>Unmeasured or incompletely controlled confounders</td>
</tr>
<tr>
<td>Problems with vaccine batch</td>
<td>Chance finding; more likely with small sample size</td>
</tr>
<tr>
<td>Waning immunity resulting in lower VE</td>
<td>Population coverage does not align with source population of cases</td>
</tr>
<tr>
<td>Different outcome or schedule is being evaluated from clinical trial</td>
<td></td>
</tr>
</tbody>
</table>

\[
\text{crude } VE = 1 - \left[ \left( \frac{PCV}{1 - PCV} \right) \left( \frac{1 - PPV}{PPV} \right) \right]
\]

VE = Vaccine Effectiveness
PCV = percentage of the cases who are vaccinated
PPV = percentage of a comparable group in the population who are vaccinated

Fig. 1. Screening Method: The relationship between % of population vaccinated (PPV), estimated vaccine effectiveness and % of cases vaccinated (PCV) (17)

Example shown in figure with red diamond is the result for vaccine with 90% estimated VE and 90% PPV.
The VE for COVID-19 vaccines can vary between products and against different outcomes \((18, 19)\). Thus, it is important to clearly define the outcome of interest and use appropriate datasets for the screening method: for example, hospitalized cases to assess VE against severe disease.

The percentage of the cases who are vaccinated (PCV in formula) can be obtained from surveillance data from national or subnational areas or from one or a few health facilities (e.g. sentinel sites). These data should be case-based to allow for stratification by vaccine target group and vaccine type. At a minimum, cases must have been eligible to receive vaccines at the time of disease onset (e.g. targeted for vaccination when vaccines were available), and case data should include symptoms (or asymptomatic infection), severity (hospitalization, severe disease or death), date of symptom onset, vaccine brand, number of vaccine doses received and date of final dose in the series. If available, variant viruses identified by genomic sequencing should be included.

We recommend the following in calculating percent of cases vaccinated (PCV) from surveillance case data:

- Only cases who were eligible for vaccination should be included in the calculation of PCV (e.g. target age group).
- Only laboratory confirmed cases should be included.
- Only cases occurring after part of the population had received the full recommended vaccination schedule (e.g. 2 doses of mRNA vaccines) should be included.
- Cases should be considered fully vaccinated if symptom onset occurred 1–2 weeks after receipt of the recommended number of vaccine doses. Cases who did not complete the recommended vaccination schedule at least 1–2 weeks before symptom onset should be excluded.
- If sufficient numbers of cases are genomically characterized (e.g., sequencing, SNP detection), cases may be stratified into vaccine reference virus or variant virus. However, one must consider the sequencing algorithm to ensure that it is randomly conducted in the population being evaluated (see 6. Analytic considerations for genomic sequencing when estimating vaccine effectiveness for new variants).
- Vaccination data should include vaccine type or brand. For vaccine type or product specific VE, only include cases completing the series of a specific vaccine type or brand in the numerator. In the denominator, include individuals eligible for vaccination with the specific vaccine and exclude cases who received other vaccine types or products or no vaccine. Because the screening method is approximate, it is most useful for estimating VE for a complete vaccine series.
- The minimum sample size will be determined by the vaccine coverage. The sample size for cases is shown in Table 1 below with a power of 80%.
- Results can also be stratified by vaccination target group (e.g. age group) and time (e.g. week or month), as has been done for influenza \((21)\).
Table 1. Minimum sample size requirements of cases for a range of vaccine coverage [Fleiss with Continuity Correction] (20)

<table>
<thead>
<tr>
<th>Expected vaccine effectiveness</th>
<th>Vaccine coverage in population being evaluated (for complete series)</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>30%</td>
<td></td>
<td>470</td>
<td>342</td>
<td>286</td>
<td>261</td>
<td>259</td>
<td>281</td>
<td>350</td>
</tr>
<tr>
<td>40%</td>
<td></td>
<td>250</td>
<td>179</td>
<td>147</td>
<td>132</td>
<td>128</td>
<td>136</td>
<td>165</td>
</tr>
<tr>
<td>50%</td>
<td></td>
<td>151</td>
<td>106</td>
<td>85</td>
<td>75</td>
<td>71</td>
<td>73</td>
<td>86</td>
</tr>
<tr>
<td>60%</td>
<td></td>
<td>98</td>
<td>67</td>
<td>53</td>
<td>45</td>
<td>42</td>
<td>42</td>
<td>48</td>
</tr>
<tr>
<td>70%</td>
<td></td>
<td>67</td>
<td>45</td>
<td>34</td>
<td>29</td>
<td>26</td>
<td>25</td>
<td>27</td>
</tr>
<tr>
<td>80%</td>
<td></td>
<td>47</td>
<td>31</td>
<td>23</td>
<td>20*</td>
<td>20*</td>
<td>20*</td>
<td>20*</td>
</tr>
<tr>
<td>90%</td>
<td></td>
<td>33</td>
<td>21</td>
<td>20*</td>
<td>20*</td>
<td>20*</td>
<td>20*</td>
<td>20*</td>
</tr>
</tbody>
</table>

*Note that a minimum sample size of cases regardless of expected VE and coverage is 20. Numbers in these cells have been changed from the calculation to reflect this.

The percentage of a comparable group in the population who are vaccinated (PPV in formula) can be obtained from administrative databases, vaccination registries or coverage surveys. Key criteria for coverage data are described below.

- Coverage data should be for the same population (e.g. city, region) and same target group in which cases were identified.

- Population coverage should only be among those who are vaccine-eligible and be stratified in the same way that the case data are stratified (e.g. by age group, by vaccine type or product and by calendar time).

- Coverage with complete vaccination series should be calculated, i.e., the percent of the source population that is fully vaccinated.

- For vaccine type- or product-specific VE, the source population should exclude persons vaccinated with other vaccine types or brands.

- Time considerations
  - Vaccination coverage in the source population 1–2 weeks prior to illness onset among cases should be used to calculate PPV to allow sufficient time to achieve full immunity.
  - For cases aggregated by month, vaccination coverage at the midpoint of the preceding month can be used.
  - If the outcome is hospitalization or severe disease, vaccination coverage 3–4 weeks prior to hospitalization, severe disease outcome or death may be used to account for time from illness onset to these more severe outcomes.

The calculation outlined above can now be populated with PPV and PCV to calculate the crude VE. 95% Confidence intervals (CI) can be calculated by using logistic regression as outlined by Farrington et al (22). An R package has been developed, which allows one to calculate the crude VE and 95% CI (available [here](#)).
Once a crude VE is calculated using the screening method, this should be compared to the vaccine efficacy or effectiveness results obtained from clinical trials or observational VE studies conducted in a similar setting, similar population and with a similar outcome (e.g. asymptomatic infection, symptomatic infection, hospitalization, severe disease, death). VE estimates can also be compared over time to detect changes. VE estimates outside the confidence intervals of prior estimates require further investigation. Low VE against a new variant virus could signal reduced protection. However, alternative explanations should also be investigated. A rigorous VE study should be undertaken if the screening method identifies a lower VE than expected.

There are several limitations to the screening method. Crude VE estimates are not fully adjusted for individual characteristics or other contributing factors (e.g. comorbidities). Biases related to surveillance (e.g. who gets tested, missing data) could impact results (23, 24). Normally, the screening method is applied to endemic diseases with stable vaccination coverage and where vaccination does not create high levels of herd immunity (25). With COVID-19, some of these attributes are unknown or may not apply. Lastly, poor quality of surveillance and coverage data that do not encompass the same population that gave rise to the cases can lead to large differences in crude VE estimates.
4. Case-only approaches to assessing vaccine effectiveness from new variants

4.1 Comparison of variant prevalence in vaccinated and unvaccinated cases

In settings with multiple SARS-CoV-2 variants identified in circulation, analyses of cases may provide early warning signals of decreased vaccine effectiveness against variant viruses. One analysis, relying solely on surveillance data, is a comparison of vaccination among cases due to variant viruses with vaccination among cases due to reference strains. While not providing estimates of VE, case-only analyses indicate whether protection against variant viruses is decreased relative to protection against vaccine-reference strains. If there is lower effectiveness against specific variants, the incidence of vaccine breakthrough cases due to variant viruses increases at a given time and place.

All COVID-19 cases included in case-only analyses must be genomically characterized. To increase sample size, SNP-detection assays can be used in contexts where variant dynamics is well characterized (e.g. L452R positive SNP detection tests may be indicative of Delta, or N501Y+E484K positive tests may be indicative of Beta and/or Gamma, if no other circulating variant carries that mutation or combination of mutations). Reference viruses may be those used in vaccine production or against which vaccines have demonstrated effectiveness, including the Alpha variant, which has become predominant in many countries (26). Cases due to reference SARS-CoV-2 viruses and variants included in case-only analyses must be identified in the same source population and during the same time period to control for changes in vaccination coverage and variant incidence.

A simple 2×2 table for case-only analysis can be set up as follows:

<table>
<thead>
<tr>
<th>COVID-19 vaccination</th>
<th>New variant case</th>
<th>Cases due to reference virus</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vaccinated</td>
<td>(A) # vaccinated new variant cases</td>
<td>(B) # vaccinated reference virus cases</td>
<td># vaccinated cases (i.e., vaccine breakthroughs)</td>
</tr>
<tr>
<td>Unvaccinated</td>
<td>(C) # unvaccinated new variant cases</td>
<td>(D) # unvaccinated reference virus cases</td>
<td># unvaccinated cases</td>
</tr>
<tr>
<td>Total</td>
<td># new variant cases</td>
<td># reference virus cases</td>
<td># genomically characterized COVID cases analysed</td>
</tr>
</tbody>
</table>

A crude odds ratio (OR) is the odds of vaccination among new variant cases (A/C) divided by the odds of vaccination among reference virus cases (B/D), or (A*D)/(B*C). OR > 1 results when odds of vaccination among new variant cases is higher than the odds of vaccination among reference virus cases. When OR is greater than 1, this indicates a reduction in vaccine effectiveness against new variants compared to effectiveness against cases due to reference viruses (27). For more advanced analysis, confounding can be adjusted for through regression models, typically logistic or log-binomial models, considering each new variant independently.

Note that the case-only analysis does not provide an estimate of absolute vaccine effectiveness (5). Although it is similar to an indirect cohort analysis for estimation of efficacy of pneumococcal vaccination (27), the indirect cohort method assumes the vaccine has no impact on the risk of non-vaccine type
pneumococci among vaccinated persons. The equivalent assumption would be that COVID-19 vaccination has no protective effect against new variants, which is likely not the case since COVID-19 vaccines so far appear to confer some effectiveness against new variants, based on existing immunological and effectiveness data (2–4, 28). Therefore, a case-only analysis can only indicate if a vaccine has a relative reduction in VE against new variants compared to reference viruses.

To adjust for potential confounding by timing of new variant circulation and vaccine rollout by age indication, stratification by calendar time should be done in a case-only analysis (if not matched by time of enrolment). Alternatively, cases due to variant viruses may be matched to cases due to vaccine-reference viruses by time of illness onset and potential confounders like age and residence (11) and vaccination odds ratios estimated as a matched analysis, using either an exact test (e.g. McNemar’s) or conditional logistic regression. Last, as noted earlier, rapid replacement, within weeks of emergence, means that the time window to undertake a case-only approach may be short, with cases due to the reference strain rapidly decreasing to levels where comparison would no longer be possible.

### 4.2 Sieve analysis

An alternative to focusing on reduced VE against defined variant lineages (i.e., VOIs/VOCs) is to ask if certain mutations or specific combinations of mutations or lineages predict reduced VE. Sieve analysis is related to the case-only analysis described above whereby sequences from vaccinated and unvaccinated cases are compared. This approach uses phylogenetic, bioinformatic and statistical methods, without requiring strong prior hypotheses on which changes in the genomic sequence will be important (e.g. cases do not need to be binned a priori). While sieve analysis was initially conceived to evaluate viral genetic features in vaccine efficacy trials comparing vaccine and placebo recipients, this approach can be applied to investigate viral determinants of reduced VE following vaccine roll out in the population. These determinants can be identified in currently circulating viral variants or to predict which future variants may be more likely to evade the vaccine immune response (29).

Sieve analysis can be used in different settings if only one variant represents almost all sequences or if there are multiple variants co-circulating. For example, Alpha represents the majority of infections in several settings, thus potentially precluding analysis of Alpha vs other viruses. However, there are sub-lineages of Alpha with specific mutations, and sieve methods can identify sub-lineages or specific mutations associated with reduced VE. Sieve analysis methods can be grouped in three categories depending on the focus of the analysis.

- Genome- or gene-based analyses characterize phylogenetic trees, viral diversity and the distance from the sequence used in the vaccine (e.g. spike protein) to evaluate whether vaccinated and unvaccinated cases differ.

- Site specific analyses are designed to test whether, for a given site, sequences among vaccine breakthrough cases are more likely to mismatch the vaccine than sequences from unvaccinated cases. A specific set of sites in the protein corresponding to the vaccines selected prior to the analysis to ensure sufficient power to detect a signal (i.e. sites that are too conserved or too variable are excluded) and additional knowledge can be used to select sites that are important for vaccine-induced immunity. These may be sites associated with ACE2 binding affinity, known antibody contact sites and sites known to affect neutralization sensitivity (e.g. recurrent mutations at specific spike sites in new variants).
Epitope specific analyses are designed to integrate multiple sites in the spike. They rely on current knowledge of the immunity induced by SARS-CoV-2 infection and vaccination, with particular emphasis on specific epitopes in the Receptor Binding Domain (RBD). Here, sites in the spike are weighted based on the strength of the interaction with specific antibodies or their potential impact on neutralization. Epitope-specific analyses are well suited to searches for evidence of potential escape from the vaccine induced response (30–32).

As with other analyses, sieve analysis requires control for confounding by time, location and demographic factors and should be done in consultation with statisticians familiar with this complex methodology.
5. Full vaccine effectiveness evaluations for new variants

If changes in COVID-19 epidemiology, examination of breakthrough cases, or case-only analyses suggest that there might be reduced VE against new variants, then a full VE evaluation might be warranted. Approaches to VE evaluations of new variants can be considered in two scenarios: a) when there is co-circulation of previously predominant variants with new variants and b) when the new variant has become the predominant strain in a population.

5.1 Evaluation of COVID-19 VE when new variant is co-circulating

A. When the majority of cases have genomic characterization

The test-negative design case-control study (TND) (5) used to estimate effectiveness of COVID-19 vaccination against disease or infection due to any SARS-CoV-2 virus may also be used to estimate COVID-19 VE against new variants. This approach is similar to a clade-specific VE evaluation for influenza vaccines (33). In this analysis, COVID-19 vaccination status (or odds of vaccination) among cases due to new variants is compared to vaccination status among persons who test-negative for SARS-CoV-2. In this TND design, the VE against cases due to new variants provides absolute, not relative, VE estimates against a specific outcome (e.g., disease, infection).

Sample size estimates for estimation of VE must be calculated for each variant and vaccine type separately, and the previously cited sample size formula/calculator can be used as described in the interim guidance (5). Variant-specific VE analyses are conducted on data from the same source population during the same time period as the COVID-19 test-negative comparison groups to control for changes in vaccination coverage.

As in the analysis of a case-control study with only one variant in circulation, the crude odds ratio (OR) for each variant in the setting of multiple variants in circulation is the odds of vaccination (i.e., # vaccinated / # unvaccinated) among cases for each variant divided by the odds of vaccination among test-negative controls. Vaccine effectiveness is estimated as \((1 – OR) \times 100\) for each variant. Note that the test-negative comparison group for analysis of COVID-19 VE against new variants will be the same group for each variant. An odds ratio can be calculated from a \(2 \times 2\) table or by logistic regression to adjust for potential confounders.

A simple table for estimation of VE for reference viruses and new variants can be set up as follows.

<table>
<thead>
<tr>
<th>SARS-CoV-2 virus</th>
<th>SARS-CoV-2 cases</th>
<th>Test-negative controls (non-case patients)</th>
<th>VE (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All SARS-CoV-2 virus (regardless of any genomic characterization)</td>
<td># vaccinated/total cases (%) vaccinated</td>
<td># vaccinated/total controls (%) controls vaccinated</td>
<td></td>
</tr>
<tr>
<td>Only genomically characterized SARS-CoV-2 viruses</td>
<td># vaccinated/total genetically characterized cases (%) vaccinated</td>
<td>Same as above</td>
<td></td>
</tr>
<tr>
<td>Reference SARS-CoV-2 virus</td>
<td># vaccinated/total reference virus cases (%) vaccinated</td>
<td>Same as above</td>
<td></td>
</tr>
<tr>
<td>New variant-1</td>
<td># vaccinated/total new variant-1 cases (%) vaccinated</td>
<td>Same as above</td>
<td></td>
</tr>
<tr>
<td>New variant-2</td>
<td># vaccinated/total new variant-2 cases (%) vaccinated</td>
<td>Same as above</td>
<td></td>
</tr>
</tbody>
</table>
A lower VE for a new variant compared to the reference virus suggests a reduced VE against that variant (particularly if 95% confidence intervals are non-overlapping.) If the VE against all SARS-CoV-2 viruses regardless of whether genomic characterization was done differs from that of all genomically characterized SARS-CoV-2 viruses, this suggests a potential bias in which viruses were selected to be sequenced.

B. When a majority of cases do not have genomic characterization

The gold standard evaluation of VE for new variants will require genomic characterization of all cases in the study population. WHO encourages genomic characterization to be as complete as possible when undertaking VE evaluations. Nonetheless, in most settings genomic characterization will not be available on most COVID-19 cases in a VE study. Only a subset of cases will have genomic characterization, or the distribution of circulating variants will be derived from population-level sampling. In these settings, the VE against new variants might only be inferred from VE evaluations. The contribution of a new variant to the overall VE found in a VE evaluation will depend on its prevalence in the population – the higher the prevalence the greater its contribution to the overall VE. However, as noted previously, a lower VE than expected in a population with a new variant in circulation could have many other causes besides a lower VE against the variant. (see Box 1).

Several considerations for interpreting VE results for new variants are applicable in VE evaluation settings without direct genomic characterization.

- There are multiple levels of uncertainty in inferring a VE for a new variant, including uncertainty of the VE estimate as well as uncertainty around the prevalence of the new variant in the test population. Prevalence of the new variant in the general population from which the sequencing data is derived (e.g. GISAID) might differ in place and time from the more localized population of the VE study.

- A qualitative, rather than quantitative, interpretation of the VE of the new variant might be more appropriate. For example, a very low or very high VE might indicate a different VE for a prevalent new variant. Studies that provide outliers in terms of VE estimates should be further investigated to discover the reasons for the lower VE. However, minor differences in the calculated VE from what is expected are likely affected by too much uncertainty and imprecision to draw conclusions about the VE for the new variant.

- Besides the new variant, the other variants in circulation might also have reduced VE compared to the referent strain for clinical trials and previous observational VE studies (e.g. previously identified VOCs). This could also account for a lower VE in addition to the new variant of interest. Of note, in observational studies Alpha has shown similar VE to that of most vaccines in clinical trials (18, 34). In the few VE studies of the other VOCs to date, Beta and Delta have shown slightly reduced VE for mild-moderate disease, particularly after one dose, compared with Alpha; there is still insufficient evidence to conclude whether Gamma has lower VE (35–41).

- Approaches to better estimate the VE for new variants include a) pooling of data from common protocols for sequenced viruses and b) looking for concordance for neutralization data and VE results for new variants, which would add credibility to the VE estimate (42, 43).

In conclusion, inferring VE for a new variant from a VE study that does not include genomic sequencing (or a surrogate, like S gene target failure for Alpha or SNP detection of mutations) is challenging and should be interpreted in the context of the prevalence of the new variant in the study population and take into account other reasons for finding a different VE than expected.
5.2 Evaluation of COVID-19 VE when new variant is the predominant circulating strain

If representative genomic sequencing has proven that a new variant has completely replaced previous viral strains, the best-practice approaches for conducting a vaccine effectiveness study are outlined in Evaluation of COVID-19 Vaccine Effectiveness: Interim Guidance (5). The outcome of interest, such as infection, symptomatic disease or severe disease should be defined a priori. Multiple study designs such as cohort and traditional and test-negative case-control studies can all be used to evaluate VE for a new variant. Evaluation of laboratory-confirmed outcomes rather than syndromic outcomes and the use of documented rather than self-reported vaccination status are still recommended. Covariates to be collected and reporting guidance as outlined in the above guidance should still be followed.

There might be some unique challenges that arise when using these same methods to evaluate new variants. If new variants are already predominant at the time of vaccine rollout, then the methods are mostly unchanged. However, if new variants fully replace previous strains by a timepoint after vaccine introduction, the country’s vaccine rollout will be at a more advanced stage when assessing VE for variants. In this case, there will be fewer unvaccinated people left in the population. It is advisable to undertake VE studies before coverage reaches 90%, because the remaining unvaccinated persons will likely have atypical infection risk, which could lead to biased VE estimates. Both cohort and case-control studies may be affected. Moreover, the testing practices in a highly vaccinated population might have changed from what they were earlier in the vaccine rollout. Contemporaneous testing practices should be evaluated at the time of a VE study for variants to assess for possible biases; this could be particularly relevant for a TND. If substantial time has elapsed since persons were vaccinated when new variants become dominant, vaccine-induced immune protection might have waned for all SARS-CoV-2 strains. Such waning could be misconstrued as lower VE against new variants. Last, the laboratory test used to identify the new variant should achieve the recommended ≥ 85% sensitivity and ≥ 98% specificity for the currently circulating variants, although this might not be known at the time of the VE study. Lower test performance for new variants can lead to misclassification, particularly in the TND. It is possible that adaptations to the currently used RT-PCR tests will be needed for future new variants.
6. Analytic considerations for genomic sequencing when estimating vaccine effectiveness for new variants

For genomic sequencing within a VE study, several additional considerations should be addressed. Viruses sequenced should be representative of the SARS-CoV-2 viruses from cases belonging to the VE study population. The selection of which viruses to sequence should be independent from vaccination status, case demographics or clinical outcomes. It will be important to sample viruses from cases matched by calendar time and by site of enrolment/source population (e.g. hospitalized vs. outpatient). Sampling can be done in several ways within a VE study (44): 1) all cases (among those whose samples are technically feasible to sequence, e.g. having sufficient viral concentration); 2) a random selection of cases within a predefined time frame; 3) a systematic random sample, such as every 4th case, depending on the total number of cases and lab capacity; or 4) the first fixed number of cases (e.g., 5 or 10), each day or week, stratified by site and age. Regardless of method, it should be ensured that each virus has the same probability of being selected. Because the sampling frame can differ by approach, we recommend consultation with a statistician to adjust for sampling fraction in analysing the data.

Several biases could affect the interpretation of VE due to non-representative sampling. For example, vaccinated individuals might have lower viral loads than those who are unvaccinated (45, 46). Viruses with Ct values > 30 are usually more difficult to sequence. This means that potentially a lower proportion of viruses can be sequenced in samples from vaccinated individuals compared to those who are unvaccinated. If viral load also varies by new variant, this could create confounding and biased VE estimates. To check for potential bias related to sequencing samples, different approaches could be taken. 1) Compare overall VE to the VE only from sequenced cases. 2) Compare demographics, site distribution, and calendar time of case with sequenced virus with cases without sequenced viruses and with test-negative controls.

As discussed, if no sequencing is performed among VE study participants, contemporaneous sequencing information obtained outside of the study can be useful to help interpret VE results in light of circulating new variants in the population. However, the population from which sequencing data is available could differ from the VE study population, and the criteria for sequencing may not be random (e.g. more vaccinated individuals targeted, specific clinical conditions targeted), leading to erroneous conclusions about the VE against new variants.
7. Bias in vaccine effectiveness studies of new variants

Any bias that can occur when estimating VE overall can occur when estimating VE for a single strain separately, including new variants (5). For example, if receiving a vaccine is positively associated with exposure to COVID-19 (e.g. because of profession, such as a health care worker), then estimated VE overall, and for each strain, may be biased downward compared to the true VE, if this exposure is not adjusted for in the analysis. If health workers are exposed to the new variants in the same proportion as others, then the relative VE should be unbiased when comparing the vaccine status of cases with different variants. Likewise, many of the other potential biases listed in Table 3 of the Interim Guidance (5) – if they apply to all variants equally – may bias absolute VE estimates but not strongly bias relative VE estimates.

Some biases might apply differently to analyses of relative VE against infections or disease with new variants. If a new variant is more likely to cause severe disease, each of the biases that apply more to severe disease will also apply more to the new variant (e.g., false negative Covid-19 test). Other, more complex, interactions may occur that lead to differential bias in estimating VE of different variants, and thus to biased comparative VE estimates. One example is the case of waning immunity over time since vaccination. If the frequency of new variant in a population is increasing with time compared to a previously predominant strain, then there will be a positive association between time of infection and chance of being infected with the variant. If either true waning of effectiveness or spurious waning due to depletion of susceptible individuals is occurring, this will produce a falsely lower estimate of VE against the new variant than the previously predominant strain, thereby potentially underestimating the VE of the new variant in a comparative VE estimate. This issue could be addressed by matching comparisons on calendar time.

If diagnostic sensitivity is lower for a new variant, this may or may not lead to biased estimates of relative VE, depending on whether the lower sensitivity is specific to vaccinated persons. As mentioned, lower viral loads in vaccinated persons combined with a suboptimal target for amplification by PCR could hypothetically make the vaccine look more effective against a new variant because of failure to see a signal in vaccinated, infected persons. If lower sensitivity is unrelated to vaccine status, this should bias absolute VE estimates but not relative, estimates.

More research is needed to better understand the biases that may arise due to possible combinations of different levels of protection from past infection with different strains, different VEs for different strains and unknown durations of protection from both post-infection and vaccine-derived immunity.
8. Conclusion

It is likely that new SARS-CoV-2 variants will continue to emerge as the COVID-19 pandemic evolves. It will be essential to assess how vaccines perform against these new variants, to inform immunization programmes, policy makers and vaccine developers. A critical piece of evidence to evaluate performance will likely come from post-introduction observational VE studies. This addendum updates WHO’s interim guidance on undertaking VE studies of COVID-19 vaccines in the setting of new variants.
References


