Q&As developed by the COVAX Regulatory Advisory Group (RAG): Oct 2020

Comparability to support manufacturing changes

Background

- Development of manufacturing processes for COVID-19 vaccines is being executed within considerably reduced timelines, and with evolving knowledge on product, analytics and process, requiring potential deferral of activities (e.g., optimization/validation) after launch.
- Compared to other modalities, vaccines are diverse products, hence the level of risks/acceptance associated to the proposals may vary depending on the prior knowledge and degree of complexity and understanding of product and process, however, general scientific principles can be agreed across product types.
- In addition, clinical and post-launch supply will require use of multiple manufacturing sites and post-approval changes to support the administration of doses to billions of patients.
- The number of batches used in the clinic (Phase 1 and Phase 3) and the urgency with which these studies are being executed result in a limited historical dataset to establish statistically-based acceptance criteria which are typically applied for comparability assessment.
- While following a manufacturing change, the question arises as to whether the post-change product is comparable to the pre-change product, to ensure that the pre- and post-change products perform equivalently. In this context, building strong, quality risk-based comparability strategies is key to support fast access to vaccinees and sustainable lifecycle management.
- Comparability approaches and burdens of proof for comparability vary greatly from country to country, as do approval timings. This can create delays in getting vaccine to many markets quickly.
- Given the challenges associated with the COVID-19 emergency, comparability assessment may be on critical path. Cross-industry reflection and engagement of Regulatory Agencies is hence of high importance, as it may provide a structured set of options to be rapidly assessed for the individual platforms/products.

Potential approaches to demonstrate comparability of COVID-19 vaccines during development and lifecycle will include

- The use of a risk-based analytical comparability assessment of manufacturing changes, for instance:
    - evaluate a subset of Critical Quality Attributes that are impacted by the proposed changes and are known (e.g., via prior/platform knowledge) to possibly have impact on safety and/or efficacy at the levels exposed to the vaccinee (when administered at the desired dose).
    - consider matrixed and bracketed approaches across DS and DP
    - assess the need of additional characterization testing to reinforce comparability data

- The use of release, forced degradation and/or characterization data to demonstrate comparability, depending on the changes being made. In addition, the comparability strategy may vary depending on the nature of the change and supporting analytical and process evolution.
- Critical quality attributes for post-change lots could be compared to lots used in the pivotal study in which clinical efficacy has been demonstrated, thereby supporting comparability based on product quality with a link to the patient without a need to obtain further clinical exposure. Assessing manufacturing variability in clinical trials and appropriate dose selection (as per discussion at 2018 EMA/FDA early access workshop) would support definition of such patient-driven acceptance criteria for comparability.
- Where prior knowledge is limited and/or in the absence of statistically-based acceptance criteria, a "clinical development"-type approach to CMC comparability may be appropriate, aimed at demonstrating the preservation of critical quality attributes without the requirement of process consistency, given the limited manufacturing history in accelerated scenarios. This is in line with ICH Q5E, stating that "the goal of the comparability exercise is to ascertain that pre- and post-change drug product is comparable in terms of quality, safety, and efficacy."
- The global use of general/broader PACMPs for routine changes (e.g. new reference standards/positive controls, new cell bank, new stock seed, changes to raw materials or excipients such as new suppliers, minor DS and DP manufacturing changes, manufacturing location or scale-up)

Approval of the original application with the comparability protocol can provide the applicant an agreed-upon plan to implement the change. Depending on the change, the applicant can provide control strategy, risk assessment, product knowledge to potentially reduce the reporting category for the CMC change.

Background materials:
- EMA FDA Stakeholder workshop on support to quality development in early access approaches, such as PRIME and Breakthrough Therapies
  https://www.ema.europa.eu/en/events/stakeholder-workshop-support-quality-development-early-access-approaches-such-prime-breakthrough
  EFPIA White Paper on CMC development, manufacture and supply of pandemic COVID-19 therapies and vaccines
  https://www.efpia.eu/media/554681/cmc-development-manufacture-and-supply-of-covid-19-therapies-and-vaccines.pdf
- CBER's Draft Guidance on Comparability Protocol
  https://www.fda.gov/regulatory-information/search-fda-guidance-documents/comparability-protocols-human-drugs-and-biologics-chemistry-manufacturing-and-controls-information

Questions:

Would the Regulatory Advisory Group agree to the following?

- Apply risk-based analytical comparability assessments of the subset of CQAs that may be impacted by the proposed changes
- In cases where prior knowledge is limited, and when there is no statistical basis for acceptance criteria due to limited number of lots, use of approaches to comparability focused on product quality expectations
- A global single approach to comparability amongst nations, considering;
  - early feedback from regulatory authorities on comparability approaches in advance of obtaining efficacy data from Phase 3 to help confirm requirements and ensure alignment on product specific approaches.
  - global use of general/ broader PACMP for routine changes/introduction of multiple manufacturing process changes, including introduction of reliance mechanisms

Answers

The Developers would need to focus on CQAs known to affect safety and efficacy and these CQAs should be well defined and supported. It is uncertain if there could be a single global approach, but the elements proposed to establish comparability seem reasonable and in line with ICH Q5E [https://www.ema.europa.eu/en/documents/scientific-guideline/ich-q-5-e-comparability-biotechnological/biological-products-step-5_en.pdf]. Moreover, the RAG supported the risk-based approach.

The only caveat is that comparability should also consider the specifics of each case. It is therefore difficult for the RAG to say in every scenario whether regulators will be able to transpose the proposed strategy to all vaccines.

It is noted that many of the developers are producing at risk and the impression is, based on what has been communicated to regulators, that a substantial amount of manufacturing data is being generated. Hence, RAG members were of the opinion that there will be sufficient manufacturing information available, which would make a risk-based approach comparability feasible. That said, it should be adequately demonstrated that lots included in a comparability exercise are reflective of lots used in clinical trials and material to be used at commercial scale. COVID-19 vaccine developers are strongly encouraged to get early feedback from regulators on their comparability approach.

Forced degradation studies are an excellent way to assess relative stability pre- and post-manufacturing change, provided the stability indicating potential of the assays is well defined. While harsh degradation conditions (e.g., oxidative and temperatures > 50°C) are reasonable initial conditions to evaluate, more appropriate conditions reflective of typical temperature excursions (e.g., ≤ 37°C) will be more biologically relevant for the evaluation of the product, assuming the stability indicating potential of the assays has already been demonstrated appropriately. It was acknowledged that there may be situations where the antigen is highly stable, but that needs to be shown with multiple orthogonal methods to provide convincing data to demonstrate such stability.

While preclinical and clinical data is important in the evaluation of stability indicating quality attributes, the high degree of variability associated with *in vivo* assays could make comparability challenging. Thus, RAG stressed that appropriately designed *in vitro* stability indicating assays can be more sensitive, robust and reproducible and are therefore preferred for quality control purposes.

RAG members stressed that there is a need for very strong analytical packages and that the analytical package must be focused on the proposed changes in the manufacturing process. Moreover, it will be important to include stability data and characterisation tests in the analytical package. If analytical

methods are changed during the development of the product, then comparability of the old and new method must be well characterised or the assessments could prove difficult. As far as possible, the analytical methods should not be modified significantly all along the clinical development phases in order to have a solid baseline for the comparability exercises.

It was pointed out that in addition to the routine release tests used in a comparability exercise, developers should consider additional characterisation tests to support comparability over the life-cycle of the vaccine. This is particularly important during the clinical development phase, up to the registration to be sure that comparability of commercial lots can be linked to batches that have been found to be safe and efficacious in clinical trials.

Clinically relevant product specification considerations: Since Phase 3 trails are generally used to demonstrate clinical consistency, there is a tendency to use lots that are relatively consistent in terms of quality attributes. This tends to lead to the establishment of narrow specification ranges, since the specifications should be linked to the clinical lots. The tighter the quality specifications are, the more likely batch rejections will be for potentially useful clinical lots. Hence, it is recommended that during early clinical development, sponsors should aim at established clinically meaningful ranges for specific CQAs. This would typically occur during phase 2a dose-finding studies to support CQAs such as potency. When correlates of protection are not defined, as is the case with COVID-19, the alternative is to perform a broader set of immunological assays (e.g. neutralising antibody titres, CMI, cytokine profile etc.) potentially on a smaller subset of subjects. Such studies should be developed in coordination with regulatory authorities.

The use of Burden of disease as end-point for efficacy

Background

The consensus for the primary efficacy objective in pivotal Phase 3 vaccine efficacy (VE) trials has been clinically symptomatic COVID-19 rather than asymptomatic SARS-CoV-2 infection. However, there is no consensus on the appropriate case definition for the primary endpoint which is reflected by the various case definitions outlined in the publicly available VE trial protocols of various developers.

It is likely that COVID-19 vaccines, like other respiratory and mucosal virus vaccines, are most effective in preventing severe disease rather than mild disease or asymptomatic infection. The vaccines will prevent disease progression and severity may shift from severe to mild (vaccine-mediated attenuated disease, VAD) among COVID-19 cases of vaccine recipients. If a vaccine acted in this way, unknown to the investigators, then, in a clinical vaccine efficacy (VE) trial with a primary endpoint of 'COVID-19 / any severity', mild COVID-19 occurring in vaccine recipients would be counted as vaccine failure (endpoint case) rather than as successful prevention of disease progression (VAD). This may be a concern when assessing 'COVID-19/any severity' as a primary efficacy endpoint.

The epidemiology of COVID-19 shows that the incidence of mild disease far exceeds severe. This makes the convincing demonstration of VE against severe disease challenging and requiring a trial size much larger than the already large size trials based on any symptomatic disease endpoint. Therefore, in these trials, severe COVID-19 is being included as a key secondary endpoint.

The analyses such as testing hypotheses or deriving confidence intervals uses the classical approach gives a score of 0 to a non-case and a score of 1 to a case. The BoD expands this approach further. It gives an integer score such as 0, 1, 2 and so on, with increasing score signifying increased severity. For example, in a VE trial of COVID-19 vaccine, a score of 0, 1, 2, 3 or 4 is given to asymptomatic infection, mild disease, moderate disease, severe disease and death, respectively. The score for each group is then, as in the classical case, the sum of the scores of the individuals. The difference in scores can be tested to calculate vaccine efficacy and a confidence interval for VE. The extension of the statistic from 0 and 1 to several positive integers makes the statistical distributions somewhat different, but not unusual.

BoD or Burden of Illness (BOI) endpoint has previously been accepted for regulatory approval of Zostavax.

The design of pivotal Phase 3 COVID-19 VE trials may benefit from including BoD as follows:

- dual primary endpoint (alongside 'COVID-19/ any severity') or
- triple primary endpoint (alongside 'COVID-19/any severity' and 'COVID-19/moderate to severe') or
- key secondary endpoint (e.g. with 'COVID-19/any severity' and 'COVID-19/moderate to severe' as dual primary endpoint)

It is unclear whether or not Phase 3 VE trials will meet the requested minimum 50% VE target (with lower bound confidence interval of >30%) against 'COVID-19/any severity' if the main effect of the vaccine is prevent severe disease with relatively lower effect on mild disease . On the other hand, Phase 3 VE trials are unlikely to be sufficiently powered to demonstrate VE against severe COVID-19 based on the targeted 150-160 confirmed cases of symptomatic disease.

A BoD endpoint seems to be an appropriate approach to assess VE against progression to severe disease and de-risk a situation in which inappropriately defined primary endpoints do not reflect an important aspect of the potential protective efficacy of COVID-19 vaccine leading to Phase 3 trial failure.

Nevertheless, pivotal Phase 3 VE trials should assess other aspects including severe COVID-19 as well as infection and transmission as additional (e.g. secondary endpoints).

Background materials

- Clinical Endpoints for evaluating efficacy in COVID-19 Vaccine Trials. Mehrotra D. et al, Ann Intern Med. 2020 Oct 22 : M20-6169. 10.7326/M20-6169
- Zostavax: EPAR scientific discussion Available at:
- https://www.ema.europa.eu/en/documents/scientific-discussion/zostavax-epar-scientific-discussion_en.pdf
- Callegaro A, Curran D, Matthews S. Burden-of-illness vaccine efficacy. Pharmaceutical Statistics 2020; 19: 636-645. https://doi.org/10.1002/pst.2020

Questions:

1) Would a BOD endpoint be acceptable as a single, dual or triple primary endpoint in Phase 3 trials to establish VE against COVID-19?

Rationale: BoD endpoints should be acceptable as separate endpoints in addition to endpoints assessing COVID-19 illness of pre-defined specific severity. This would de-risk pivotal Phase 3 VE trials in case other endpoints do not meet pre-specified criteria because of low VE (COVID-19 of any severity with mostly mild cases) or because of insufficient number of cases (e.g. severe COVID-19, hospitalization, death).

2) If a BoD endpoint was to be an acceptable primary endpoint, what would be the requested success criteria for a BOD endpoint?

Rationale: The success criteria for VE as recommended by FDA as well as draft WHO PQ/EUL guidelines is a point estimate that is ≥50% with a lower bound confidence interval of >30%.

Answer:

The RAG fully understands the rationale behind this question, i.e. vaccines may show better efficacy in preventing severe disease than mild disease. This is at least what is known for other respiratory viruses. It was therefore recognised that there is risk that an all-comer study using "any severity" as the primary endpoint could fail but, that the secondary end point of "severe disease" could be statistically significant. To accept this approach would take some more discussion as it is not straight forward. One challenge would be how to transform the outcome into a meaningful indication.

RAG members had diverging viewpoints. Some were open to discuss the possibility of using two primary end-points: any severity" and "severe disease spectrum". This can be tackled methodologically and statistically in ways that are acceptable to regulators. However, it is a prerequisite to have a good definition of what is meant by "moderate" and "severe disease". There are concerns that definitions currently used in some studies are not acceptable. COVID-19 vaccine developers should make an attempt to have homogeneous definitions of moderate and sever disease. A simple way would be to just follow what WHO has published in terms of clinical management of patients with COVID-19 [https://www.who.int/publications/i/item/clinical-management-of-covid-19].

Other opinions were that a BoD endpoint which evaluates severity adjusted vaccine efficacy, could provide useful data on the target population who might benefit most from the vaccination. However, as it is based on a scoring system instead of a binary outcome of "yes" or "no" to the presence of symptomatic disease, it could potentially bias the results in favour of the vaccine (i.e. inflate the vaccine efficacy) in the scenario where the vaccine only attenuates the disease, but does not actually reduce the overall incidence of the disease. Therefore, the preference was for the primary efficacy endpoint to focus on a reduction in the incidence of symptomatic disease, which would provide an unbiased estimation of the true vaccine efficacy. However, BoD could be considered as a key secondary endpoint, in addition to the existing conventional secondary endpoints that already measure severe disease.

Some RAG members were concerned that adequate data collection for a BoD endpoint could be challenging in large studies.