Taylor & Francis
Taylor & Francis Group

Check for updates

Challenges in Assessing the Impact of the COVID-19 Pandemic on the Integrity and Interpretability of Clinical Trials

Mouna Akacha[a], Janice Branson[a], Frank Bretz[a,b], Bharani Dharan[c], Paul Gallo[c], Insa Gathmann[a], Robert Hemmings[d], Julie Jones[a], Dong Xi[c], and Emmanuel Zuber[a]

[a]Clinical Development & Analytics, Novartis Pharma, Basel, Switzerland; [b]Section for Medical Statistics, Medical University of Vienna, Vienna, Austria; [c]Clinical Development & Analytics, Novartis Pharmaceuticals, East Hanover, NJ; [d]CONSILIUM Salmonson & Hemmings, London, UK

ABSTRACT

The COVID-19 pandemic has a global impact on the conduct of clinical trials of medical products. This article discusses implications of the COVID-19 pandemic on clinical research methodology aspects and provides points to consider to assess and mitigate the risk of seriously compromising the integrity and interpretability of clinical trials. The information in this article will support discussions that need to occur cross-functionally on an ongoing basis to "integrate all available knowledge from the ethical, the medical, and the methodological perspective into decision making." This article aims at facilitating: (i) risk assessments of the impact of the pandemic on trial integrity and interpretability; (ii) identification of the relevant data and information related to the impact of the pandemic on the trial that needs to be collected; (iii) short-term decision making impacting ongoing trial operations; (iv) ongoing monitoring of the trial conduct until completion, including the possible involvement of data monitoring committees, and adequately documenting all measures taken to secure trial integrity throughout and after the pandemic, and (v) proper analysis and interpretation of the eventual interim or final trial data.

1. Introduction

The novel coronavirus (SARS-CoV-2) is a new strain of coronavirus that had not previously been identified in humans. The coronavirus family is known to cause illness in humans, from common cold to more severe or even fatal diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). On 30 January 2020, the World Health Organization (WHO) declared the outbreak a public health emergency of international concern. On 11 March 2020, WHO characterized COVID-19 as a pandemic.

The COVID-19 pandemic has a global impact on the conduct of clinical trials of medical products. Challenges may arise, for example, from quarantines, site closures, travel limitations, interruptions to the supply chain for the investigational product, or other considerations if site personnel or subjects become infected with COVID-19. More specifically,

- healthcare systems in many regions of the world are overloaded, and medical personnel is drawn into COVID-19 related activities;
- government restrictions put in place have direct consequences on trial monitoring, data collection and drug supply activities;
- subjects may be reluctant to attend scheduled visits (hesitance to go to places where people may be ill, to travel, to be in contact in the public, etc.);
- subjects may be unable to attend scheduled visits due to

  - health issues (e.g., hospitalization or death) related to COVID-19; or
  - government restrictions (e.g., self-isolation, quarantine, lockdowns);

- subjects may need to take additional medications to treat COVID-19 symptoms, or other health issues related to, for example, the confinement situation (such as antidepressant drugs) which were not anticipated at the design stage of a given clinical trial.

These challenges may lead to difficulties in meeting protocol-specified procedures, including administering or using the investigational product or adhering to protocol-mandated visits, efficacy assessments, and laboratory/diagnostic testing. More specifically, these complications can result in

- compromised trial data (e.g., missed assessments, trial procedures changed to virtual assessments) and
- challenges in the interpretation of clinical trial results due to systematic biases that can be introduced in a number of ways (e.g., subjects with certain co-morbidities or in certain age classes are under-represented).

The extent of the challenges will depend on, for example, the duration of the current COVID-19 pandemic, the number of impacted subjects, the disease condition being studied, and various trial design elements (e.g., trial duration, visit schedule

intervals). In some cases, the disruption may be so substantial that continuation of the trial may not be viable. In other cases, trials may be able to continue; however, it seems likely that data collected beyond the current point in time will be suspect, posing a number of methodological challenges (Fleming, Labriola, and Wittes 2020; McDermott and Newman 2020; Meyer et al. 2020; Wolkewitz and Puljak 2020). By "suspect" we mean that the data will be different in nature, collected under different circumstances, in a different manner, and may be questionable in terms of its quality.

In recognition of the extraordinary situations the pandemic presents, the U.S. Food and Drug Administration (FDA 2020) and the European Medicines Agency (EMA 2020a) have released guidance for sponsors on how they should adjust the management of clinical trials and participants during the COVID-19 pandemic. In essence, both guidelines sensibly ask sponsors to collect the reasons for protocol deviations (PDs) and discontinuation of clinical trial elements (e.g., investigational treatment) related to COVID-19 in all ongoing trials.

EMA also published a second guidance (EMA 2020b) on actions that sponsors should take for ongoing clinical trials affected by the COVID-19 pandemic. This document builds upon EMA (2020a) and provides points to consider to mitigate the risk of seriously compromising the integrity and interpretability of ongoing trials due to COVID-19 while safeguarding the safety of trial participants as a first priority.

The importance of collecting relevant data related to the pandemic is not to be underestimated. Absence of such information can put data integrity, trial integrity and clinical trial interpretability at risk and likely impact future submissions. Data integrity is defined as the extent to which all trial data are complete, consistent, accurate, trustworthy, and reliable throughout the data lifecycle (WHO 2019). Trial integrity is a broader concept relating to trial conduct more broadly, which encompasses data integrity and which refers to the ability of a trial to produce results that can be relied on for decision making. In particular, this means that results are not affected by (unknown) biases. For example, unblinding during an ongoing trial can result in a loss of trial integrity. Moreover, cohort effects and informative dropout mechanisms if unknown and not adequately accounted for can lead to a loss of trial integrity. The extent to which data and trial integrity are affected has an impact on clinical trial interpretability and the conclusions that can be drawn from the data collected.

In this article, we provide

- a categorization of complications due to the COVID-19 pandemic, which will likely have statistical and trial integrity implications and thus may lead to updates of trial protocols and statistical analysis plans of ongoing trials;
- a discussion on unforeseen intercurrent events (ICH 2019) and their impact on original trial objectives;
- a flowchart that illustrates some of these considerations and possible actions, depending on whether a given trial is halted, continued or stopped due to COVID-19;
- suggested points to consider to support cross-functional discussions that need to occur on an ongoing basis in view of the complications and subsequent implications mentioned above;

- a rationale for collecting detailed information on the complications that arise due to the pandemic; and
- a discussion of statistical considerations including missing data, consistency of treatment effects and implications on trial power if a trial was to stop early because of COVID-19, accompanied by an R Shiny app to facilitate power consideration.

2. Intercurrent Events, Estimands, and Missing Data

The COVID-19 pandemic results in various complications for subjects and sites participating in ongoing clinical studies. These complications have an impact on various clinical trial aspects: trial conduct, data integrity, the scientific question that can be addressed by the trial as well as the statistical analyses. Generally, the complicating events due to COVID-19 can be categorized into those that are more of an administrative or operational nature, and those that are more directly related to the effect of COVID-19 on the health status of subjects. Note that some of these complications may be a direct consequence of measures taken because of the pandemic.

2.1. Category 1: Complications Due to Administrative or Operational Challenges

Complications falling into this category include:

- treatment discontinuation due to drug supply issues;
- treatment discontinuation due to subject concerns;
- inability to perform important procedures (e.g., biopsies, laboratory/diagnostic tests);
- missed visits (e.g., subject preferences, self-isolation or government restrictions such as quarantines or lockdowns);
- visits outside of the designated time window;
- altered or compromised visits due to overloads of health system (e.g., remote communication with sites rather than in-person, different site or provider, or local vs. central review).

2.2. Category 2: Complications Directly Related to the Impact of COVID-19 or the Pandemic on the Health Status of the Subject

Complication falling into this category includes:

- treatment discontinuation due to COVID-19 symptoms;
- intake of additional medications to treat COVID-19 symptoms;
- intake of anti-inflammatory treatments when being COVID-19 infected (even though the interplay with COVID-19 may not be yet well-understood);
- death due to COVID-19;
- inability of COVID-19 infected subjects to attend scheduled visits;
- health issues induced or exacerbated by the government restrictions or the health system overload (e.g., neglecting of underlying chronic conditions, worsening or newly occurring depression, negative impact on quality of life, etc.).

Common to all of these complicating events is that they were not foreseen at the design stage of the ongoing trials. In randomized blinded controlled clinical trials, these complications may be expected to apply similarly to different treatment arms. However, this may not be the case in open label trials, where differences in complications may be expected between the treatment arms due lack of blinding. Likewise, randomization may not protect us in settings where efficacy or safety are moderated in one treatment arm in a way that changes the possibility to contract the virus or the outcome of a COVID-19 infection, for example, one of the treatment arms contains a drug with an immunosuppressive mode of action. In this case, we may expect to see more treatment discontinuations or missed visits due to subjects' concerns in the treatment arm with the immunosuppressive mode of action. The impact of the complications will likely vary across different regions and sites, even within the same country. It will also vary depending on characteristics of the actual subjects. For example, it is known that the elderly and those with underlying conditions such as asthma, diabetes, etc. are at higher risk of missing visits and adverse consequences from COVID-19 (see, e.g., Brown et al. 2020; Cawthon et al. 2020).

Some of these complications lead to unforeseen intercurrent events in the sense that they affect either the interpretation or the existence of the measurements associated with the clinical question of interest (ICH 2019) while others prevent relevant data being collected and result in a missing data problem. In the following, we will discuss both aspects in turn.

2.3. Intercurrent Events and Estimands

It is important to distinguish between COVID-19 pandemic related and unrelated intercurrent events, for example, "treatment discontinuation due to drug supply issues caused by the pandemic" versus "treatment discontinuation due to lack of efficacy." Relevant intercurrent events that are not related to the pandemic were probably already foreseen at the design stage of the study. In contrast, intercurrent events related to the pandemic, for example, death due to COVID-19 or treatment discontinuation due to pandemic related drug supply issues, were neither foreseen nor addressed at the design stage.

Changes in the clinical question of interest with respect to intercurrent events foreseen at the trial design stage may be controversial and should be duly justified. However, there is a need to articulate the question of interest in respect of unforeseen intercurrent events due to the pandemic. There is also a benefit to this, and the consequent revision of the estimand, for the purpose of discussion between stakeholders and alignment of methods for data handling and statistical analysis to the clinical question of interest.

Unforeseen intercurrent events can lead to various challenges and questions. For example, are observations in the dataset representative of the original population of interest or do the data predominantly reflect subjects without certain co-morbidities, or only younger subjects? Some of the unforeseen intercurrent events may even result in the need to change certain endpoints, if the ones originally specified cannot be collected or assessed as planned. As mentioned above, some subjects may also need to pause or discontinue their investigational treatments. This has a direct impact on collected efficacy and safety data and begs the question whether the strategy originally identified for the question of interest in respect of treatment adherence is still the relevant one? More specifically, the question of interest in respect of "treatment discontinuation" in the original estimand might now need to be rephrased to account for treatment discontinuations related to the COVID-19 pandemic and treatment discontinuations unrelated to COVID, or into more categories reflecting the need to employ a different strategy for each. Specific challenges may result in the setting of open-label studies due to lack of blinding, where differential rates of unforeseen intercurrent events such as treatment discontinuations might be observed dependent on the nature of the treatments.

For most settings, it is conceivable that the original trial objective and treatment effect of interest will remain unchanged. However, the unforeseen intercurrent events due to COVID-19 introduce ambiguity to the original research question and teams need to discuss how to account for them (Akacha et al. 2017; Akacha, Bretz, and Ruberg 2017).

No adaptation of the original estimand implicitly suggests a treatment policy approach for all unforeseen intercurrent events. Such a strategy may be relevant in certain settings, for example, when only few unforeseen intercurrent events occur. In general, however, it seems plausible to frame clinical questions in the presence of the unforeseen intercurrent events in category 1 using a hypothetical estimand strategy. That said, different hypothetical strategies could be considered (ICH 2019). For example, does interest lie in the treatment effect in the absence of COVID-19, that is, in a world where the disease does not exist? Alternatively, are we interested in the effect of the treatment in a world where individuals can suffer from COVID-19 infections but in the absence of the administrative and operational challenges caused through the pandemic? It is conceivable that medical practice may be slightly different in the future as a consequence of the current pandemic even if everyone is vaccinated against the virus.

The complications listed in category 2 are directly related to the health status of the subjects and the role of a hypothetical strategy is therefore less clear. This holds in particular when changes in the health status due to the pandemic are related to the studied condition or treatment. For example, how should we account for deaths due to COVID-19 in a lung cancer outcome trial where death is an outcome of interest? Should this be counted as an additional event for the outcome of interest or addressed with an appropriate choice of strategy as an intercurrent event? The impact of such intercurrent events—which may occur at different rates in the different treatment arms—needs to be carefully assessed on a case-by-case basis.

If interest lies in a hypothetical strategy for any of the unforeseen intercurrent events, then the interplay between foreseen and unforeseen intercurrent events should also be carefully considered. For example, if interest lies in the treatment effect applicable to a world where COVID-19 does not exist, then subjects would not suffer from the unforeseen intercurrent events, however, they are at risk of experiencing the foreseen intercurrent events, for example, treatment discontinuation due to an adverse event. This needs to be taken into account when predicting plausible hypothetical trajectories for subjects.

More generally, assumptions for the predictions of trajectories need to be aligned with the hypothetical strategy of interest. In this context, it appears helpful to ask "from where/whom do we borrow information to predict the hypothetical measurements of interest?" For example, do we borrow information from within a given trial or do we leverage information from external sources? Additional considerations include whether sufficient data and information is available to inform a prediction model and the need to adequately account for the uncertainty in the predictions. While this discussion may be similar in spirit to that of "missing data" problems, the challenge with predicting hypothetical trajectories is conceptually different. In the latter, we may well have collected data after an intercurrent event, for example, measurements after treatment discontinuation, however, this data is deemed not relevant for the scientific question of interest. Instead, hypothetical trajectories that are more aligned with the estimand of interest are predicted.

2.4. Missing Data

For the discussion of missing data challenges due to the COVID-19 complications, it is worth revisiting the definition of missing data according to the ICH E9 addendum (ICH 2019): "Data that would be meaningful for the analysis of a given estimand but were not collected. They should be distinguished from data that do not exist or data that are not considered meaningful because of an intercurrent event." It is of importance that missing data relates to data that would be meaningful if it had been collected. Missing data clearly leads to a loss of information, but it can also introduce selection bias. Therefore, appropriate missing data handling approaches aligned with the estimand of interest and plausible sensitivity analyses need to be specified.

Several of the complications that are listed in category 1 result in missing data, for example, the inability to perform important procedures like biopsies during the pandemic or government restrictions which prevent subjects to attend scheduled visits. These reasons for missing data are likely of external nature, that is, they are probably unrelated to the health status of subjects or the treatments under investigations. In such cases, assuming an ignorable missingness process seems plausible. By ignorable missingness we mean that a missing at random (MAR) process holds and that valid inference can be based on the available data only, that is, there is no need to model the missingness process (Molenberghs and Verbeke 2007).

In contrast, missing data as a result of the complications listed in category 2 may be due to the health status of the subjects and not ignorable in nature. Consider a chronic obstructive pulmonary disease trial where a study participant is in the intensive care unit to treat severe respiratory symptoms caused by COVID-19 and therefore cannot attend a scheduled visit. In this case, the missingness process depends on unobserved health status information which is relevant for the disease and treatment under study.

The plausibility of the assumed missingness process needs to be duly justified and the robustness of conclusions to the assumptions made needs to be assessed through sensitivity analysis. Given the potentially large amount of missing data, the focus for sensitivity analyses should lie on plausible assumptions and not on overly "conservative" assumptions. In this context, any approach chosen to deal with missing data needs to adequately account for the added uncertainty, for example, through the use of multiple imputation rather than single imputation approaches.

The aspects around intercurrent events, estimands, and missing data which were discussed in this section need to be carefully considered within clinical teams and may require different measures to be taken in the course of the trial conduct, including the monitoring of those events, and/or that protocols and statistical analysis plans are updated, potentially in agreement with regulatory agencies.

Statistical considerations beyond those presented here for missing data and the prediction of hypothetical trajectories are provided in Section 4. A flowchart illustrating some of the high-level considerations provided in this section is given in Figure 1.

3. Need to Collect Detailed Information on COVID-19 Related Complications

An immediate challenge facing clinical trials teams is the need to decide how best to capture the additional information needed to enable an assessment of trial integrity. For example, do we collect information that allows us to identify the complications listed in Section 2 and for how long they lasted? Moreover, are we collecting whether deviations such as treatment discontinuations or interruptions are due to direct or indirect COVID-19 reasons and can we distinguish these from deviations that are unrelated to the COVID-19 pandemic? The importance of this information was also highlighted in the regulatory guidelines that were referenced in Section 1. Many of the situations we are now facing were not foreseen at the start of ongoing trials and our standard data capture methods were not set up to deal with them. In the following, we will share some considerations on what information needs to be collected and how to implement the collection of the additional data in practice.

3.1. What to Collect?

The complications listed in category 2 relating to a subject experiencing a confirmed or suspected COVID-19 infection are perhaps the simplest to record. Clinical trials routinely record data on adverse events during the course of the trial, and the changes in study medications, use of concomitant medications and the outcomes associated with those.

Other complications due to the pandemic may apply at three levels. On a subject level, subjects may discontinue treatment earlier than planned, or withdraw their consent and discontinue from the study. For subjects remaining in a trial, changes may occur at a visit level, for example, some visits may be postponed beyond the protocol allowed time window or missed treatment may be temporarily interrupted. The complications may also apply at an assessment level, that is, some assessments may not be able to be done even if a visit takes place, or may be performed differently, perhaps remotely. Ideally in all of these cases it would be valuable to know what had happened and why, however, a pragmatic approach may focus attention on the subject and visit level and aim to capture changes at the assessment level for
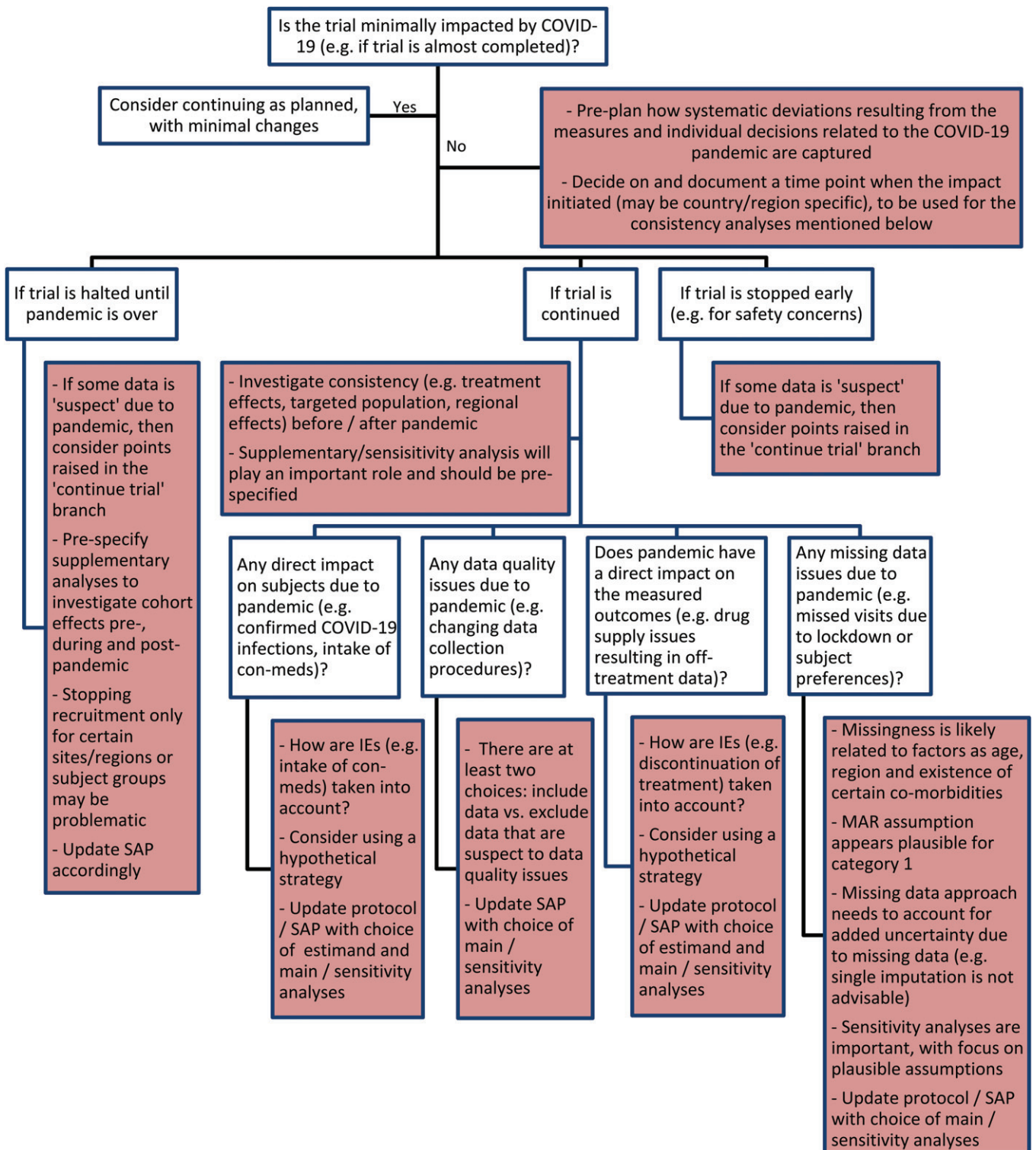
Figure 1. Flowchart illustrating some of the considerations provided in Section 2, together with possible actions implied by those considerations. Note that this is not a decision tree. Certain aspects (e.g., safety, operational challenges) are not covered. IE: intercurrent event; MAR: missing at random.

only those assessments that relate to primary or key secondary objectives.

Regulatory guidelines (EMA 2020a; FDA 2020) recommend to collect data on the complications due to the COVID-19 pandemic, for example, the reasons for missing data, and to assess the impact of these complications on the robustness of

conclusions that can be drawn from the data. A key consideration is how much detail on the specific reason for the change should be captured. The collection of more detailed reasons than "due to COVID-19 pandemic" can be very helpful to assess the occurrence of unforeseen yet important intercurrent events and should include specifics such as

- due to the subject's own (confirmed or suspected) COVID-19 infection,
- due to general quarantine,
- due to site specific issues,
- due to lack of study drug availability,
- due to subject concerns.

This information will aid the phrasing of clinical questions of interest in respect of additional intercurrent events, see the discussion in Section 2. This information will also be crucial for informing plausible missing data assumptions for primary and sensitivity analyses, and as such may impact the data handling strategy for the key analyses.

If we can capture these differences in reasons, we will hopefully be in a good position to understand what happened in the trial and draw sensible conclusions about the impact the pandemic had on the trial and its ability to answer its objectives.

3.2. How to Collect?

As mentioned earlier, data related to a subject's COVID-19 infection is most easily collected via standard practices for recording adverse events and concomitant medications, although it may be necessary to provide more detailed guidance to sites on the need to record suspected as well as confirmed cases. Supplemental questionnaires or electronic case report forms (eCRF) may be used to record further details such as dates and outcomes of tests for the virus.

For other complications, in practice, methods for data collection may come down to a choice between modifying the eCRF or collecting data via some other means. Modifying the eCRF involves changes on the database side, programming of new edit checks around the new pages, translation into different languages, release, training and roll-out to teams and investigational sites. This can create a large burden, both from the data management end and on the site side. Teams should give careful consideration as to whether adding additional data entry work to sites already on the frontline of running trials is a feasible approach.

Alternatives to changing the eCRF may be considered and FDA (2020) indicates that sponsors may develop processes that enable systematic capture of such information. One method is to adapt existing PD processes. Typically, identification of potential deviations occurs more on the sponsor side with sites responding to queries where needed. This may be a more feasible approach for reducing the burden on sites. Additionally, processes and systems for capturing PDs are often in place and can be more readily adapted to capture these deviations and the reasons for them. A concern sometimes raised about this strategy is that it will lead to a large increase in PDs. However, regulatory guidelines indicate that an increase in PDs during the pandemic is to be expected, and that PDs will be assessed in a proportionate manner (EMA 2020a). One approach could also be to summarize the important PDs that were already defined at study start separately from the new COVID-19 specific PDs that were added during the conduct of the study.

In addition to the aforementioned strategies, conducting telephone calls to collect information on the complications listed in Section 2 could be a viable and complementary approach.

4. Additional Statistical Considerations

In Section 2, we discussed statistical challenges around missing data and the prediction of hypothetical trajectories that results from the pandemic. However, the COVID-19 pandemic leads to various additional complications that deserve a discussion. For example, single arm trials that were aiming at contrasting results of historical data to the data collected in the trial may need to reassess the level of comparability and relevance of the available historical data. Data quality concerns may need to be assessed through sensitivity analyses.

While there are numerous statistical challenges as a result of the pandemic, we will in this Section focus on aspects around consistency of treatment effects and power implications. Some of these high-level considerations are summarized in a flowchart which is given in Figure 1.

4.1. Consistency

The consistency of treatment effects by region or of the population before, during, and after the COVID-19 pandemic may be questioned. Sometimes, consistency of results pre- and post-pandemic might not be expected, for example, when study treatment intake is interrupted during the COVID-19 pandemic due to drug supply issues. It is likely that data before the COVID-19 pandemic will receive larger focus, as this is the only data certain not to be subject to COVID-19 related influences, whether known or unknown. Investigations to address consistency (e.g., tests of interaction over time), may receive additional importance, but it is not clear what type of consistency analyses are useful and what their operating characteristics are. For example, Friede and Henderson (2009) investigated tests for heterogeneity in the context of adaptive designs to assess whether treatment effect estimates differ significantly before and after an interim analysis. Also, Kunz et al. (2020) discussed the use adaptive designs to combine information across stages (e.g., pre-/during/post-pandemic) and/or to allow for unplanned mid-trial modifications to respond to the pandemic. The use of these and other methods will benefit from additional scientific discussions and regulatory guidance. Regardless of the specific analytical approach, the dates determining the pre-/during/post-pandemic periods should be prespecified and documented properly, together with the associated choice of consistency analyses.

4.2. Power Implications

It may be helpful in making certain decisions to approximate a trial's current operating characteristics based on data existing so far (existing in the field, even if not yet collected), and thus not compromised by COVID-19 concerns. As described previously, power implications may arise for various reasons: a trial may not be feasible to continue because of COVID-19 issues; or it may be possible to continue but with a lesser amount of information than planned; or, even if a trial can continue, perhaps the data beyond some point in time may later be viewed

by regulatory agencies as compromised, so that main attention will be given to data already existing by that point. As a simple generic illustrative example: consider a trial designed for 90% power to detect a specified clinically relevant treatment effect; in an analysis using data where only 60% of subjects reach the trial's endpoint, how much does this reduce power?

Certainly, the original methodology used for study design can be extended to address power for any amount of data. However, as a broadly applicable illustration, consider a normally distributed test statistic for a primary endpoint, in a study designed for level α and power $1 - \beta$, requiring N subjects. Say that we only obtain a fraction p of the targeted information (e.g., pN subjects or events). The resulting power can be shown to equal

$$P\left(Z > \left(1 - \sqrt{p}\right) z_\alpha - \sqrt{p} z_\beta\right),$$

where z_γ denotes the $(1 - \gamma)$-quantile of the standard normal distribution. Thus, for example, say that we designed a study for 90% power but only obtained 60% of the targeted subjects. The power for the same magnitude of effect would be

$$P\left(Z > \left(1 - \sqrt{0.6}\right) \cdot 1.96 - \sqrt{0.6} \cdot 1.28\right)$$
$$= P\left(Z > -0.55\right) = 0.71.$$

Note that we did not need to enter the hypothesized treatment difference or standard deviation, as the power can be determined solely through the level, original power, and amount of data relative to the original design. In addition, the same calculation can be applied to binary outcomes, whether the data is analyzed on the scale of difference of means or log odds ratio, and for time-to-event outcomes where the proportion of information is the ratio of the observed number of events to the number planned in trial design. Finally, we note that in many situations, the "effective" current sample size may be higher than the current number of subjects or events, because partial data on subjects may be informative for those subjects' eventual outcomes, perhaps incorporated through a modeling approach or multiple imputation analysis.

We have developed an R Shiny app that aims to facilitate a quick simple consideration and quantification of this issue. It is quite broad and generic. It can be applied to any design for a two-arm comparison with a normal or asymptotically normal endpoint; this includes continuous, binary, time-to-event outcomes, and potentially others. All that needs to be specified is the significance level and power specified at the trial design stage, and the amount of statistical information available for analysis (e.g., the fraction of subjects with data for a continuous endpoint, or the proportion of subjects with events relative to the design plan for a time-to-event outcome). Other trial parameters, such as hypothesized effect size estimates, nuisance parameters (e.g., variance), randomization ratio, etc., need not be specified, and are therefore assumed not to be affected. This power calculation was independently also developed by Kunz et al. (2020). With our agreement, the app now also incorporates various extensions, for example, to group sequential designs as described in Friede and Henderson (2009).

The app can be accessed at *https://power-implications. shinyapps.io/prod*. It has different tabs that display information related to this question in different manners. A full description

of how to use the app and interpret its results, and the methodology underlying the calculations, appears in the help tab of the app.

5. Discussion

The COVID-19 pandemic has a global impact on the conduct of clinical trials of medical products. To assess the impact on trial integrity, we generally need to think about the consequences of the pandemic on various interrelated aspects of a clinical study:

- Trial conduct and feasibility, for example, can the population of interest be recruited and can the treatment be administered?
- Data integrity, that is, can the data of interest be collected in a complete and reliable manner?
- Scientific objective(s) and estimand(s), for example, what is the impact of the additional intercurrent events on the population, endpoint and treatment of interest?
- Statistical analyses, for example, can statistical analyses deliver reliable results targeting an agreed estimand based on the currently available data and the potential to collect further data considering operational restrictions and data integrity?

Ultimately, trialists will need to determine whether there is feasibility and value in continuing ongoing clinical trials and how best to amend protocols and analysis plans. In the online appendix, we provide an extensive list of concrete questions which aim at helping in the decision process for the impact assessment.

In terms of the scientific objective and estimands, we recommend to distinguish between COVID-19 pandemic related and unrelated intercurrent events. For pandemic-unrelated foreseen intercurrent events there may be no need for action as changes in estimand regarding foreseen intercurrent events may be controversial and need to be duly justified. For pandemic-related intercurrent events unforeseen at the design stage, we recommend to consider revising the existing estimand definition. A hypothetical estimand strategy seems to be plausible for intercurrent events related to operational challenges, although the hypothetical scenario needs to be well described, as discussed in Section 2. The relevance and acceptability of a hypothetical estimand strategy seems to be less clear for intercurrent events related to health status, for example, death due to COVID-19 in a cardiovascular outcome trial where death is an outcome of interest. Importantly, it is critical that we can reliably estimate the targeted outcome with justifiable and plausible assumptions. Reliable statistical inference on the targeted estimand is important. If this seems not possible (e.g., limited understanding of disease or drug to impute/predict missing or hypothetical data), an alternative estimand should be chosen. Ultimately, the appropriate choice of suitable estimands (and aligned analysis approaches) requires cross-functional discussions across all stakeholders (sponsors, regulatory agencies, etc.).

There are various additional considerations which we did not cover in this article, but which would benefit from additional dialogue between industry, regulatory agencies and academic partners. For example, some trial results might fail to reach

formal statistical significance due to a smaller amount of data for reasons such as we have described previously. To compensate for information lost because of the pandemic, it might be envisaged to increase sample sizes or to extend follow-up times in time-to-event trials. An alternative approach is to treat beyond the primary time point so that assessments can be made once site visits are again possible to validate assessments made remotely at the primary time point, or to facilitate modeling of outcomes that would have been observed had it been possible to take a measurement at the primary time point. Other approaches to compensate for lost information can be considered as well, such as leveraging data on short-term endpoint(s) that are correlated with the primary response or to integrate data from external sources, to augment the control arm, or pool trial data (Hemmings 2020).

As much remains to be learned on the COVID-19 disease manifestations, treatments, and pandemic distribution, it appears that monitoring the status of ongoing trials will play an important role. However, it may take particular care to ensure that this is implemented in a manner that protects the integrity of trial conduct. The use of data monitoring committees (DMCs) has been advocated to this effect by regulators (EMA 2020b), and certainly DMCs will have a role to play in reviewing relevant summaries and analyses. In the interest of trial integrity and interpretability of final results, study sponsors (or other trial management personnel, e.g., a steering committee) should consider possible modifications to study design and conduct, and extension of the analysis plan, based on sound scientific and statistical rationale (which might lead, e.g., to changed sample size or follow-up duration, modified group sequential or futility schemes, additional safety outcomes or subgroups to be monitored, etc.). This may broaden the scope of potential recommendations that the DMC is charged to consider as they perform their monitoring function.

The COVID-19 pandemic and the related measures have also the potential to impact the types, incidence, severity and duration of AEs collected for investigational treatments and in control groups, for example, when subjects take concomitant medications to treat COVID-19 symptoms or when subjects that do not want to leave the house refrain from seeking medical advice when suffering from side effects (Nilsson et al. 2020). As in the case of efficacy data, the complications and intercurrent events listed in Section 2 are also relevant when considering safety data, for example, they may lead to some under- or over-reporting of some events. In the case of efficacy data, we discussed that a hypothetical strategy may address a question of interest for the intercurrent events that result from administrative or operational reasons. It remains to be discussed whether this is equally plausible and acceptable for safety data and how we can we derive consistent benefit-risk conclusions from a trial if the volume of intercurrent events is large.

Supplementary Materials

In the online appendix, we provide a list of concrete questions that may help in the decision process for the impact assessment, such as whether to stop, halt or continue the trial, with or without changes to the protocol and trial conduct.

References

Akacha, M., Bretz, F., Ohlssen, D., Rosenkranz, G., and Schmidli, H. (2017), "Estimands and Their Role in Clinical Trials," *Statistics in Biopharmaceutical Research*, 9, 268–271. [421]

Akacha, M., Bretz, F., and Ruberg, S. (2017), "Estimands in Clinical Trials—Broadening the Perspective," *Statistics in Medicine*, 36, 5–19. [421]

Brown, E. E., Kumar, S., Rajji, T. K., Pollock, B. G., and Mulsant, B. H. (2020), "Anticipating and Mitigating the Impact of COVID-19 Pandemic on Alzheimer's Disease and Related Dementias," *The American Journal of Geriatric Psychiatry*, DOI: 10.1016/j.jagp.2020.04.010. [421]

Cawthon, P., Orwoll, E., Ensrud, K., Cauley, J. A., Kritchevsky, S. B., Cummings, S. R., and Newman, A. (2020), "Assessing the Impact of the COVID-19 Pandemic and Accompanying Mitigation Efforts on Older Adults," *The Journals of Gerontology*, Series A, DOI: 10.1093/gerona/glaa099. [421]

EMA (2020a), "Guidance on the Management of Clinical Trials During the COVID-19 (Coronavirus) Pandemic," available at *https://www.ema.europa.eu/*. [420,423,424]

——— (2020b), "Points to Consider on Implications of Coronavirus Disease (COVID-19) on Methodological Aspects of Ongoing Clinical Trials," available at *https://www.ema.europa.eu/*. [420,426]

FDA (2020), "Guidance on Conduct of Clinical Trials of Medical Products During COVID-19 Pandemic Guidance for Industry, Investigators, and Institutional Review Boards," available at *www.fda.gov*. [420,423,424]

Fleming, T. R., Labriola, D., and Wittes, J. (2020), "Conducting Clinical Research During the COVID-19 Pandemic: Protecting Scientific Integrity," *Journal of the American Medical Association*, DOI: 10.1001/jama.2020.9286. [420]

Friede, T., and Henderson, R. (2009), "Exploring Changes in Treatment Effects Across Design Stages in Adaptive Trials," *Pharmaceutical Statistics*, 8, 62–72. [424,425]

Hemmings, R. (2020), "Under a Black Cloud Glimpsing a Silver Lining," *Statistics in Biopharmaceutical Research* (under review). [426]

ICH (2019), "Topic E9(R1) on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials," available at *www.ich.org*. [420,421,422]

Kunz, C. U., Joergens, S., Bretz, F., Stallard, N., Van Lancker, K., Xi, D., Zohar, S., Gerlinger, C., and Friede, T. (2020), "Clinical Trials Impacted by the COVID-19 Pandemic: Adaptive Designs to the Rescue?," *Statistics in Biopharmaceutical Research* (under review). [424,425]

McDermott, M. M., and Newman, A. B. (2020), "Preserving Clinical Trial Integrity During the Coronavirus Pandemic," *Journal of the American Medical Association*, 323, 2135–2136. [420]

Meyer, R. D., Ratitch, B., Wolbers, M., Marchenko, O., Quan, H., Li, D., Fletcher, C., Li, X., Wright, D., Shentu, Y., Englert, S., Shen, W., Dey, J., Liu, T., Zhou, M., Bohidar, N., Zhao, P.-L., and Hale, M. (2020), "Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic," *Statistics in Biopharmaceutical Research* (in press). [420]

Molenberghs, G., and Verbeke, G. (2007), *Missing Data in Clinical Studies*, Chichester: Wiley. [422]

Nilsson, M. E., Crowe, B. J., Anglin, G., Ball, G., Munsaka, M., Shahin, S., and Wang, W. (2020), "Clinical Trial Drug Safety Assessment for Studies and Submissions Impacted by COVID-19," *Statistics in Biopharmaceutical Research* (under review). [426]

WHO (2019), "Guideline on Data Integrity—Working Document QAS/19.819," available at *http://www.who.int/*. [420]

Wolkewitz, M., and Puljak, L. (2020), "Methodological Challenges of Analysing COVID-19 Data During the Pandemic," *BMC Medical Research Methodology*, 20, 81. [420]