



---

# PHUSE Clinical Data Scientists Guide to Studies Impacted by COVID-19

## Table of Contents

Introduction.....	1
Regulatory Background.....	1
Patient Disposition/Discontinuation .....	1
Concomitant Medications/Procedures .....	2
Exposure/Drug Accountability .....	2
Protocol Deviations.....	4
Missing Data.....	5
Efficacy.....	11
Adverse Events .....	13
Laboratory Tests .....	14
Conclusion.....	17
Disclaimer .....	17
References .....	18
Acknowledgements .....	19



## Introduction

As the COVID-19 pandemic severely impacts many facets of human activity around the world, the pharmaceutical industry is being presented with significant challenges related to clinical trial research and development activities. Regulatory authorities have released guidance documents focusing on the impacts to study start-up activities, changes to ongoing study procedures, and items considered urgent safety matters during this pandemic.

This guide considers the impacts of COVID-19 on the collection and analysis of clinical trial data, offering guidance from industry experts on what clinical data scientists [1] (i.e. those who analyse data collected in clinical trials) can expect in the short term relating to the studies impacted by the COVID-19 pandemic, as well as consideration of future implications. It is important that the reader consider the specific circumstances of each impacted trial including subject population, indication being studied, length of trial, number of subjects, and local restrictions. Many sponsor companies have formed COVID-19 task force teams, which are publishing documents to supplement company SOPs and best practices. The reader is advised to consult with your specific sponsor company's rulebooks and instruction documents. The collection and analysis of data in trials for the treatment of COVID-19 is out of scope of this guide.

The statements from the US Food and Drug Administration (FDA) "Ensuring the safety of trial participant is paramount" [2] and the European Medicines Agency (EMA) "The ability to confirm eligibility and to conduct key safety assessments and trial evaluations is of particular importance" [3] both highlight the importance of safety monitoring of ongoing studies during the COVID-19 pandemic. How sponsors achieve this will differ, and different trials may require different approaches depending on the indication or the mode of action of the investigational or comparator drug(s). A number of considerations are raised in both the FDA guidance [2] and the EMA guidance [3] relating to the conduct of the study. These include, but are not limited to, the conversion of physical visits into phone or video visits or utilising alternative locations for assessments, changes to the study visit schedule, pausing trial recruitment or withdrawing subjects from trials. Any steps taken by a clinical study sponsor relating to the operational aspects of a clinical trial in order to guarantee the safety and wellbeing of patients may have an impact on the safety analysis of the study. Clinical data scientists will need to consider the actions taken by the sponsor when preparing to perform these analyses.

It is best practice to focus on the statistical analysis plan when designing clinical studies and in defining how and what data is collected. This ensures the necessary data is collected and helps to mitigate the risk of excess data collection and the need to update data collection design during study conduct to address analysis needs [4]. While most studies impacted by the COVID-19 pandemic will be in the study conduct phase, the same principles should be applied to any updates made to the data collection design. It will be important that any updates made to the operational aspects of the study and to the data collection design are properly reflected in protocol amendments.

## Regulatory Background

In the current, constantly evolving, atmosphere sponsors have been faced with urgent decisions regarding each clinical study in their portfolios, regardless of development stage. Challenging decisions based on mitigating the risks to study participants are being made on whether to delay initiation of new clinical studies, temporarily halt or terminate ongoing clinical studies, or update study protocols as necessary. These decisions must be based on the current regulatory guidance issued relevant to the COVID-19 pandemic, including those published by the FDA [2], the EMA [3] and many other countries (including individual member states within the EU) such as the UK [5]. This guide is intended to supplement the information in these guidance documents and is in no way meant to replace or contradict instructions from the health authorities.

Given the scope and breadth of the guidance documents, as well as the multiple changes required to be considered for any given study, it is essential to collaborate with all relevant team members to ensure all updates are managed appropriately. It will also be beneficial to discuss proposed approaches with health authorities and individual review divisions as part of taking decisions for individual scenarios.

## Patient Disposition/Discontinuation

During the COVID-19 pandemic study participants may additionally discontinue from the study or study treatment either due to a COVID-19 infection or to the impact of COVID-19 (e.g. travel restrictions, quarantine or site closure). This additional data may need to be collected and analysed to establish the impact of COVID-19 on a clinical study or program.

### DATA COLLECTION

Depending on the study design, clinical data scientists may see changes or updates to the collection of patient disposition data including enrolment, treatment disposition, study disposition, and epoch disposition. Where additional data is required, sponsors may find it more appropriate to add additional CRF pages as opposed to amending the existing CRF to minimise the updates to already collected data. Changes or updates to each disposition type are summarised below:

To establish whether the primary reason for discontinuation is due to COVID-19, sponsors may choose to take a number of different approaches. These will be driven by circumstances at the sponsor company and should also take site burden into account. Some sponsors have provided instructions to sites on how to complete the disposition forms (e.g. instructions to add the prefix "COVID-19:" and to specify details such as illness, quarantine and incorporated control measures). Alternatively, some sponsors may choose to add an additional question to establish whether the reason for discontinuation was related to COVID-19. Finally, some sponsors may add additional codelist items to existing discontinuation questions.

### DATA TRANSFORMATION

If a free text approach to data collection is followed, it may be

necessary to manually review data. Programming should be cautious in scanning the free text for flagging (e.g. “Not related to COVID-19” should not be flagged as a COVID-related event). The non-standard variable (NSV) DS.DSEPRELI [6] should be added to the dataset to indicate that the disposition reason is related to COVID-19.

### DATA ANALYSIS

It is important to understand that the COVID-19 pandemic is not the featured discussion and for updates to the analysis to be kept minimal. Study discontinuation/treatment discontinuation/delayed reason can be analysed following the below scenarios:

The basis for disposition analysis is described in the PHUSE White Paper “Analyses and Displays Associated with Demographics, Disposition, and Medications” [7]. Where the impact of COVID-19 on the study is considered minimal, Nilsson et al. [8] recommend that the standard disposition tables are sufficient for reporting data as collected. A supporting listing detailing the reason for discontinuation could be used to identify COVID-19 related discontinuation. Some sponsors may choose to either add an additional column to display the number of reasons for discontinuation related to COVID-19 or to present additional rows for the reasons for discontinuation related to COVID-19.

Where the impact of COVID-19 on the study is considered more significant, sponsors may consider adding a separate table for Deaths due to COVID-19 as part of the COVID-19 impact analysis. If the CSR has a special section for COVID-19 impact tables, then sponsors may decide whether to add a separate summary table of subjects who discontinued due to COVID-19.

## Concomitant Medications/Procedures

For ongoing studies, if a subject has a COVID-19 infection, the concomitant and procedures administered for treating the infection are collected as concomitant medications and procedures.

### DATA COLLECTION

In many studies, the existing collection of concomitant medications will be sufficient to capture the data required to analyse concomitant medications. Some sponsors may instruct sites to use set text to populate fields relating to the indication for the concomitant medications. Alternatively, some sponsors have chosen to add specific questions to the data collection either as additions to existing pages or as new CRF pages. These may relate to whether the medication was administered due to a COVID-19 infection or, for particular concomitant medications/procedures (e.g. oxygen therapy or ventilation), more details relating to the administration. If additional data is collected, then it should be mapped to the appropriate SDTM domain. These will include CM, PR and potentially the device domains using the variable SPDEVID to link the data.

### DATA ANALYSIS

The basis for concomitant medication analysis is described in the PHUSE White Paper “Analyses and Displays Associated with Demographics, Disposition, and Medications” [7]. For

the majority of studies it will not generally be warranted to make changes to standard analyses or to produce additional analyses [8]. Where a substantial number of subjects receive treatment for a COVID-19 infection, some sponsors may feel that is necessary to perform analyses to establish if there is an imbalance between concomitant medications among treatments for COVID-19 infections that would be important when reviewing adverse event summaries.

## Exposure/Drug Accountability

The concept of exposure to a study drug within protocol-defined windows is critical to all interventional clinical trials. The COVID-19 pandemic has affected the ability of clinical trial participants to attend sites and has interrupted the supply chain for drug products. The safety of clinical trial participants is the primary concern in conducting clinical trials, but the effects on the analysis from missed dosing or early discontinuation of study drug and impact on accountability from alternative sourcing of intervention must also be accounted for.

The FDA guidance [2] and the EMA guidance [3] on the conduct/management of clinical trials during the pandemic address issues that may arise due to interruptions to the protocol planned drug exposure and drug accountability during trials. These considerations include whether it is in the best interest of a clinical trial participant to continue drug exposure per protocol, or for the participant to discontinue study drug exposure, or end their participation in the trial entirely.

Any steps taken by a clinical study sponsor may have an impact on the safety analysis of the study. Steps taken that result in changes to the number of participants dosed, or the amount of overall time spent by participants on the drug, may affect the analysable safety population of the study. Clinical data scientists will need to consider the actions taken by the sponsor when preparing to perform these analyses.

### DATA COLLECTION AND MAPPING TO SDTM

Analysis needs should determine whether additional data will be collected, the types of data that need to be collected, and flagging records for COVID-19 relationship, as proposed in the “Guidance for Ongoing Studies Disrupted by COVID-19 Pandemic Version 1.0” [6]. The CDISC guidance will detail how things are collected, but analysis should dictate what is collected.

When determining what data should be collected, it is important for sponsors to consider that outside of trials in which COVID-19 is the indication, COVID-19 data is not the main topic of collection. Sponsors should take care not to collect data for which an analysis need has not been identified.

If analysis needs determine that COVID-19 related data is to be collected, then a sponsor’s data entry guidelines should clearly state at what timepoints the data should be collected, and when the data should be reviewed. Sponsors should also consider data consistency across all CRFs for a trial as the COVID-19 impact will be the same for all data points. Historical data may need to be collected – for example, if a trial participant missed a dose because a site was closed, but the sponsor had not already updated data collection to note site closures related to COVID-19, it may be appropriate to query the site to determine

whether the missed dose is considered COVID-19 related. If data collection methods for a particular CRF page cannot be easily changed, then other mechanisms should be evaluated to collect this data, such as updates to the protocol deviation data collection process.

As detailed in the CDISC “Guidance for Ongoing Studies Disrupted by COVID-19 Pandemic Version 1.0” [6], the variables below are SDTM variables that might be used to identify COVID-19 impacted records in the EC, EX and DA domains.

EX and EC domains:

- --ADJ
- --RSDISC

Non-Standard Variable in the SUPPEX

Variable	Label	Type	Codelist	Role	Origin
EXRSINT	Reason for interruption	text		Non-standard Record Qualifier	CRF
EXEPADJI	Epi/Pandemic Related Adjustment Reas Ind	text	NY	Non-standard Record Qualifier	CRF
EXEPINTI	Epi/Pandemic Related Interrupt Reas Ind	text	NY	Non-standard Record Qualifier	CRF
EXEPDSCI	Epi/Pandemic Related Discontin Reas Ind	text	NY	Non-standard Record Qualifier	CRF

Non-Standard Variable in the SUPPEC

Variable	Label	Type	Codelist	Role	Origin
ECREASOC	Reason for Occur Value	text		Non-standard Record Qualifier	CRF
ECEPADJI	Epi/Pandemic Related Adjustment Reas Ind	text	NY	Non-standard Record Qualifier	CRF
ECEPDSCI	/Pandemic Related Discontin Reas Ind	text	NY	Non-standard Record Qualifier	CRF

Studies where exposure records as collected contain missed visits or missed doses should follow SDTM implementation guidance to include an EC domain to contain this data. The EX domain is not intended to contain records of interventions which did not occur, only interventions which did occur.

An example of how a team may choose to capture study treatment interruption is by adding the below to the Control Terminology list for ECADJ (EC Reason for dose adjustment): SUBJECT RELATED IMPACT; SITE RELATED IMPACT; SUBJECT DIAGNOSED WITH COVID-19 INFECTION.

DA domain:

Non-Standard Variable in the SUPPDA

Variable	Label	Type	Codelist	Role	Origin
DACNTMOD	Contact Mode	text	CNTMODE	Non-standard Record Qualifier	CRF
DAEPCHGI	Epi/Pandemic Related Change Indicato	text	NY	Non-standard Record Qualifier	CRF

### DATA MONITORING

Due to the pandemic, both the FDA [1] and EMA [2] have highlighted that different approaches to monitor both the completeness and quality of data collected as part of clinical trials may be impacted. It is the responsibility of the sponsor to ensure that this is carefully monitored to ensure patient safety. This may include the “optimizing of central or remote monitoring programs to maintain oversight of clinical sites” [2]. Much of the impact of COVID-19 on a clinical study will be identified by protocol deviations and protocol amendments. Missing data and protocol deviations will be discussed in separate sections of this document, but these will potentially be important to establish whether this impact is balanced across treatment arms.

Additional data monitoring considerations should be put in place if other CRF forms related to the study medication are built into your trial. For example, if you have a form to collect the status of study medication at the end of the trial in addition to the study medication CRF form, then you must monitor the data collection in both CRF forms for consistency.

Sponsors conducting interim analyses to be included in regulatory submissions may need to account for incomplete data verification and explain when data verification will be completed, particularly for studies where the protocol is not amended, to reduce the percentage of source data verification required.

During the COVID-19 era, drug accountability may pose special data monitoring and collection challenges for Investigational Products (IPs) that are normally administered in a healthcare setting. The FDA states: “In all cases, existing regulatory requirements for maintaining investigational product accountability remain and should be addressed and documented” [2]. Thus, teams will need to carefully assess how they can monitor actual IP use, as well as give ample thought as to how COVID-19 may affect IP storage conditions. For



example, if the protocol indicates pharmacy dispensing for self-administration at home, and this is changed to direct-to-patient shipments, then a protocol amendment would be required to permit home delivery of investigational products [2].

## DATA ANALYSIS

Sponsors should collect data for the ongoing study to support the analysis. Clinical data scientists (clinical and statisticians in particular) should be driving the discussion on what will be included in the Clinical Study Report (CSR) and how newly collected COVID-19 data will impact it. Impacts on investigational product that affect analysis include:

- missed dose, delayed dose, incorrect dose
- dose obtained outside of the study
- non-centrally dispensed dose.

Consideration should be given to categorising these occurrences, as related to COVID-19, to allow for analysis flexibility. For example, it may be important to understand and analyse missed doses due to non-COVID-19 related site closure vs. COVID-19 related site closure. It may also be important to determine whether significant number of dose changes occurred, thus impacting analysis populations or sub-populations. If this occurs, the team will need to carefully evaluate the potential need for protocol amendments.

Studies may also deem it important to drill down to actual COVID-19 status, with categories such as Confirmed COVID-19, Probable for COVID-19, Negative for COVID-19. The number of impacted subjects and types of analyses to be performed will drive whether this level of detail is needed.

Finally, the impact on time-dependent variables and outputs needs to be evaluated. Each study should review displays to determine whether algorithm updates are necessary for variables such as time on treatment, duration of exposure, time since last dose, and time to event.

## Protocol Deviations

Protocol Deviations are a key source of insight regarding the conduct of our trials related to ensuring patient safety and the integrity of our trial data. The details of definitions of Protocol Deviations may differ among sponsors. Below we propose classification from one of our member companies:

A Protocol Deviation is any non-adherence to the protocol-defined study conduct for a specific subject.

Major Protocol Deviations are any deviation which has the potential to impact or impacts subjects' rights, safety or well-being, or the integrity and/or result of the clinical study. Typically, Major Protocol Deviations are predefined on a Protocol Deviation criteria list.

Potential Major Protocol Deviations are the deviations not listed on a predefined criteria list but that may meet criteria of a Major Protocol Deviation. Such potential Major Protocol Deviations require prompt review and documented confirmation of their classification.

The last category is Minor Protocol Deviations, which simply indicates a deviation not classified as major.

The data associated with these deviations is therefore an integral part in creating an overall picture that describes the impact of COVID-19 and the measures taken to manage and mitigate its effects within our trials. This section considers the impacts of COVID-19 on the collection and analysis of Protocol Deviations and offers guidance from industry experts on what clinical data scientists can expect in the short term, as well as consideration of future implications.

## DATA COLLECTION

While the pandemic's impact on the ability to conduct clinical trials is reflected in the rise of Protocol Deviations, all Protocol Deviations and issues, related and not related to COVID-19 impact, need to be documented and reviewed per your company's standard procedures. Study teams may take this opportunity to review and possibly update their predefined criteria of Major and Potentially Major Protocol Deviations in light of the pandemic.

To easily identify Protocol Deviations related to the pandemic, it is recommended site staff use a standard text string such as "COVID-19" at the beginning of the deviation description. Consistent labelling and adherence to the standard text string will also enable programmatic identification of pandemic-related records. In addition, monitors and site staff should ensure Protocol Deviations descriptions include the specific reason of relatedness [2] to COVID-19.

Examples of Protocol Deviations related to COVID-19 impact are missed doses, impact to study procedures, out of window or missed subject visits, missed disease evaluation, visits using alternative modalities, etc. Study teams can consider doing a retrospective review of previously documented Protocol Deviations and/or issues for COVID-19 relatedness and, if appropriate, amend to add a prefix description as identified above. No new CRFs or modification to existing CRFs is recommended.

## DATA MONITORING

The COVID-19 pandemic has created the situation where, due to Shelter-in-Place and travel restrictions, monitors may face delays in visiting sites to gather additional details required, perform Source Data Verification (SDV) and/or Source Data Review (SDR). This may impact timely sponsor oversight of the study.

To help with oversight and future data reporting, proper identification of COVID-19 related Protocol Deviations, a certain naming convention can be used. For example, COVID-19 related Protocol Deviations could be marked using a prefix "COVID" or "COVID-related", as mentioned above. This approach will help with future analytics and distinction between "standard" Protocol Deviations (which would occur during "normal" study conduct) versus pandemic-related.

Although it is rather difficult to fully replace a clinical research associate's work at the site in identifying Protocol Deviations, a robust central monitoring can offer substantial support in oversight. For example, certain central monitoring techniques

may be helpful in identifying data patterns indicating Potential Protocol Deviations to further investigate. Also, using central analytics properly coded Protocol Deviations could show patterns, distinction by priority or trending by deviation type for further analysis.

In scope of study conduct and following regulators' recommendations, changes in protocol conduct necessary to assure patient safety can be immediately implemented with companies, ensuring the subsequent review by the Independent Review Board (IRB) and notification to the FDA [2]. Also, companies must follow the rule that any necessary modification to protocol-specified procedures that occur prior to IRB approval and submission to the FDA must be reported as a Protocol Deviation. This approach allows for necessary flexibility when change needs to be implemented fast, but also ensures proper tracking of situation-driven introduced deviations.

### DATA ANALYSIS

While it is the responsibility of monitors to appropriately capture and document all COVID-19 related Protocol Deviations, review should take place to confirm relatedness. Clinical data scientists should work collaboratively across disciplines and within study teams to ensure any analysis variables needed to identify COVID-19 related Protocol Deviations and associated reasons follow CDISC guidelines (i.e. COVID-19 flag variable, Protocol Deviation reason variable).

If it is the company practice to only report major deviations, then it is recommended to explore the minor deviations for any textual mention of anything relating to COVID-19. It may also be necessary for minor deviations that refer to COVID-19 to be reported as major deviations, particularly if the deviation is relating to missing visits or assessments related to primary efficacy or safety parameters. It may be important to document the reasons for missing visits for future regulatory review. If the method of data capture does not support this, then alternative methods of collecting this data, such as Protocol Deviations, should be considered.

A potential approach for reviewing and reporting Protocol Deviations starts first with a manual review of the deviation descriptions at which time COVID-19 relatedness classification can be confirmed/updated. Among this subset, the associated reasons can also be identified and classified into the following 6 categories:

- Confirmed epidemic/pandemic infection
- Suspected epidemic/pandemic infection
- Subject movement restricted due to epidemic/pandemic infection
- IMP shipping delayed/blocked due to epidemic/pandemic infection
- Sponsor action due to epidemic/pandemic infection
- Site action due to epidemic/pandemic infection.

This work occurs after data collection but before data transformation (SDTM). In many cases, sponsors will need to report summaries relating to the number of Protocol Deviations due to COVID-19 and the specific reasons for the deviations to help evaluate the impact of the pandemic on the study.

## Missing Data

As noted in recent trial conduct guidance published by the US FDA, "Changes in study visit schedules, missed visits or patient discontinuations may lead to missing information (e.g. for protocol-specified procedures). It will be important to capture specific information in the case report form that explains the basis of the missing data, including the relationship to COVID-19 for missing protocol-specified information (e.g. from missed study visits or study discontinuations due to COVID-19). This information, summarised in the clinical study report, will be helpful to the sponsor and FDA" [2]. The following bullets describe some expected scenarios of missing data during the COVID-19 pandemic:

- Patient early discontinuation might cause the value of primary outcome to be missing.
- Missed visits might cause missing assessments, no procedure performed, no lab sample collected, etc.
- Drug supply delay might cause delayed or missed doses.
- Patient reported outcomes might not be consistently recorded due to variant reasons related to the pandemic, including cancelled/delayed visits.
- Paper diaries might not be returned to the site for data entry in time for Data Monitoring Committee, or other decision-making.
- Permanent data missing due to unresolved queries or unverified data.
- Site closure or site staff not available for data entry, even though patients visited the site.

### PREVENTION OF MISSING DATA AND MITIGATIONS TO CONSIDER

Mitigating the impacts of missing data is important in every trial. The pandemic, however, has the potential to increase the amount of missing data. As described in the EMA Guidance on the Management of Clinical Trials During the COVID-19 Pandemic, "Various challenges exist which result in restrictions of visits to healthcare facilities, increased demands on the health service and changes to trial staff availability. Trial participants may also be required to self-isolate, which can make it difficult for investigators to maintain their medical oversight. These challenges could have an impact on the conduct of trials, such as the completion of trial assessments, completion of trial visits and the provision of Investigational Medicinal Products" [3]. As such, it is critical to carefully plan and implement strategies to minimise the likelihood of missing data as much as possible. However, any approaches must guarantee the safety of the participants involved.

Strategies to address missing data have been widely documented. We list a few of those study conduct and data collection strategies below which are particularly relevant during the pandemic:

- Reduce required visits and amount of data collected.
- Expand visit windows.
- Adopt data collection methods that don't require face-to-face visits; use video visits.
- Utilise home nursing when participants are unable to travel to the site.
- Support the use of local laboratories.
- Contact participants via phone or use telemedicine.
- Enhance participant contact; keep participants engaged in the

study with incentives, visit reminders, phone calls to monitor status.

- Changes in data collection methods might result in a protocol amendment.

The mitigation of missing data in clinical trials is a team effort. All members of the clinical trial team involved in the design and execution of a clinical trial have a role to play in increasing retention and reducing missing data [9]. A clinical data scientist is no exception. While many mitigation strategies to reduce the volume of missing data are operational and a focus of the clinical trial site staff, the clinical data scientist must also carefully understand and evaluate the mitigation strategies being considered.

The clinical data scientist should provide input on the data collection strategies and feasibility of using the collected data for statistical analysis and reporting. The clinical data scientist has an important role to educate and inform the clinical trial team members on the use of the clinical data for downstream analysis needs. As an example, the clinical data scientist should evaluate the impact of using data collected from a local lab instead of a central lab. The clinical data scientist should ensure data collected using other means such as telemedicine for AE reporting are properly collected in the clinical trial database, keeping in mind the ability to utilise the data and, appropriately, for statistical analysis and reporting programs. (For example, verbatim and non-standard text reported in a comment field or bubble may be difficult to programmatically consume and therefore be ineffective for statistical analyses and reporting.)

Besides providing input to the design and execution of a clinical trial to increase retention and reduce missing data, the clinical data scientist must remain flexible as new data presentations may be needed in-life to evaluate missingness and identify any potential shortcomings of the available data. The talents of a clinical data scientist to handle, manipulate, and synthesise the collected data in order to monitor the trial, build knowledge and inform the clinical trial team about the degree of missingness is critically important.

### MONITORING MISSING DATA FOR ONGOING TRIALS

With all reasons of missing data in mind, the methods of review of clinical trial data must be examined and mitigating actions put into place. This might mean the adjustment of existing processes or it might mean that the review of clinical data should be addressed in a new alternative way.

### PROTOCOL DEVIATIONS

As discussed in the Protocol Deviations section, the collection of Protocol Deviations should be reviewed. If it is the company practice to only report major deviations, then it is recommended to explore the minor deviations for any textual mention of anything relating to COVID-19, bearing in mind that there could be many versions of text used. It is advised that site staff are advised to use some standard text, e.g. COVID-19, at the beginning of any deviation description text. However, one company that utilised this noted 14 ways of this being written, including COVID – 19, or \*COVID-19, so this must be explicitly followed. It is also recommended that any minor deviations that refer to COVID-19 should be reported as major deviations, particularly if the deviation is relating to missing visits or

assessments related to primary efficacy or safety parameters. It may be important to document the reasons for missing visits for future regulatory review. If the method of data capture does not support this, then alternative methods of collecting this data, such as Protocol Deviations, should be considered.

### CRITICAL VARIABLES

Teams need to check that critical variables are complete and that they are Source Data Verified and Source Data Reviewed to a level that the study team feel the endpoint will not be negatively affected with regards to the study data interpretation.

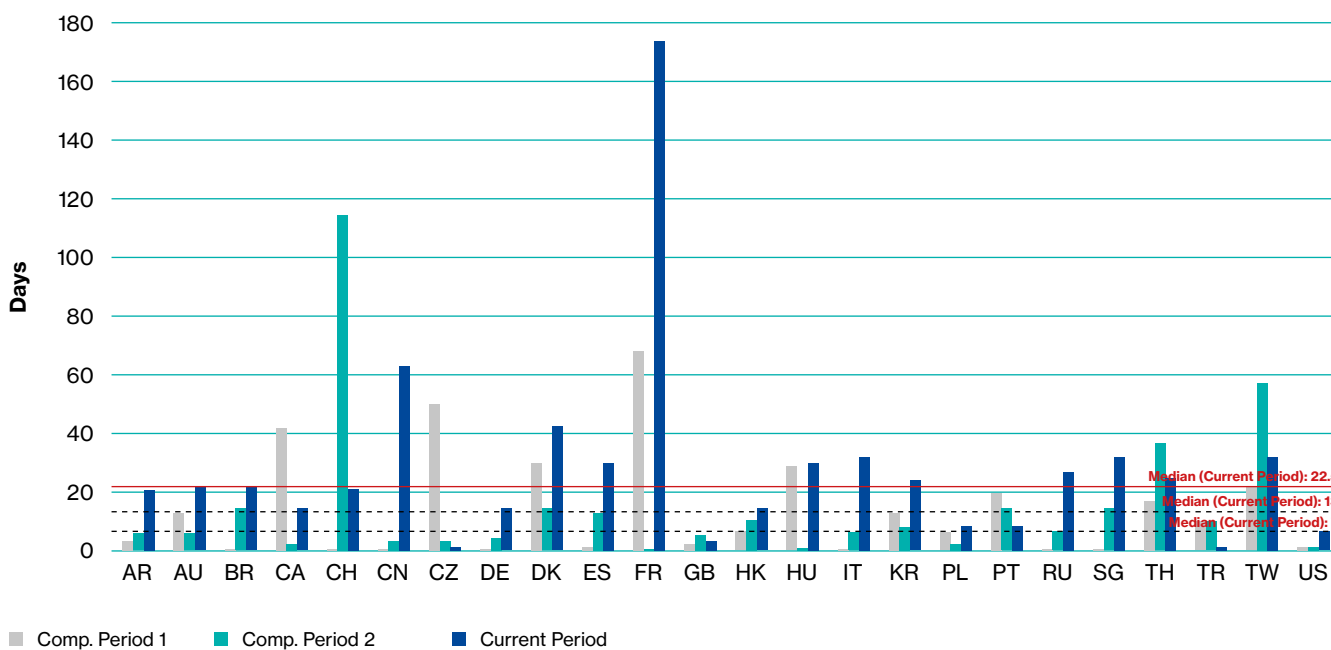
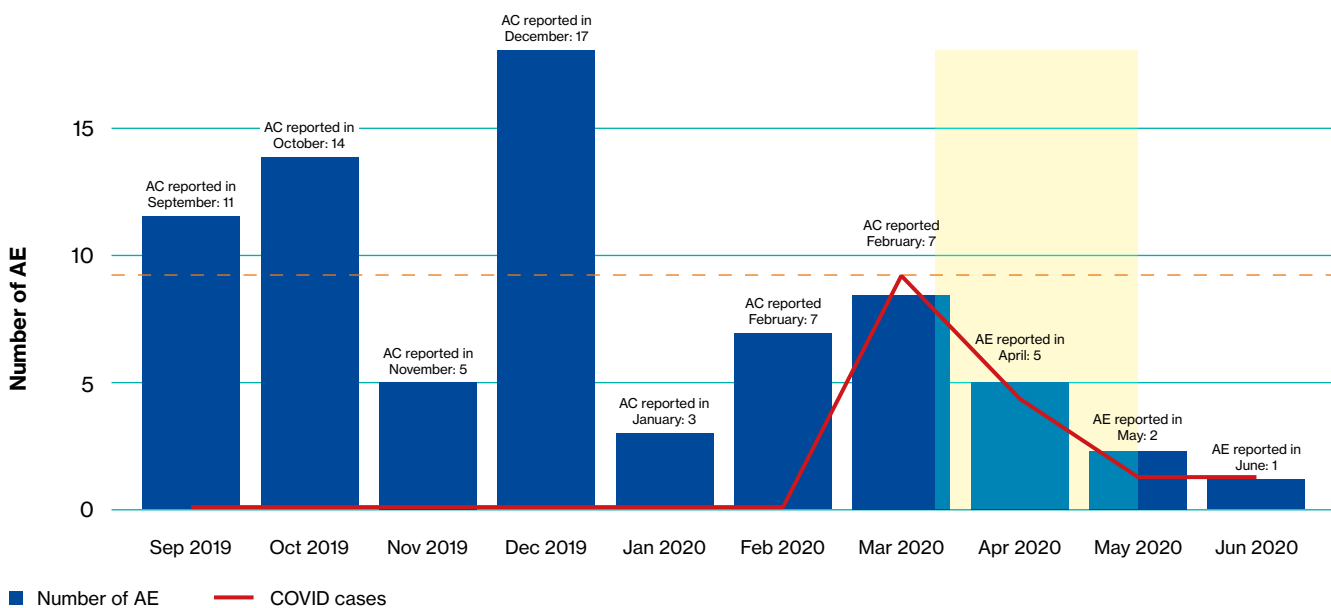
Critical variables of the utmost importance should be notified to the site. If the risk impact to these variables is high, then it may be necessary to consider delaying any data cuts or database locks and this in turn may lead to a review/amendment of the protocol. Further, any risk assessment tools would need to be enhanced to include a new risk and mitigating method to account for COVID-19.

It is recommended to put Remote Monitoring Guidance into place, where law permits (remote data review (RDR conducted by various functional departments) and remote site contact, e.g. telephone calls, email conducted by CRAs. Rather than updating functional plans (e.g. Trial Monitoring Plans, Statistical Analysis Plans), it could be advised to create functional plan addendums. This would then mean that once the pandemic is over, the original functional plan can be returned to use, and the addendum filed in the trial master file.

It is proving to be very useful to visualisations and/or reports to help study teams in their review of the amount and percent of missing visits, missing events or critical assessments. Innovative visualisations highlight the impact and missingness over time and enable the reviewer in real time to monitor the extent of the impact. Further, we can investigate the incidence (rate) of adverse event or protocol deviation reporting and the association with the COVID status in each country by utilising the data from the WHO website. This enables insights into how the event reporting has been affected by the lockdown in individual countries and how quickly the recovery of reporting can be seen.



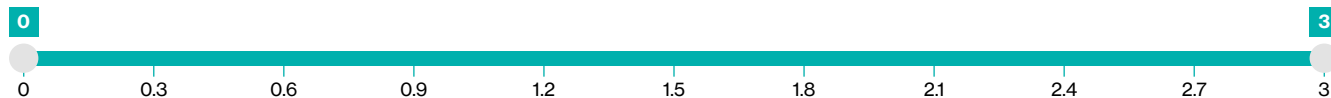
AE reporting over months and COVID evolution



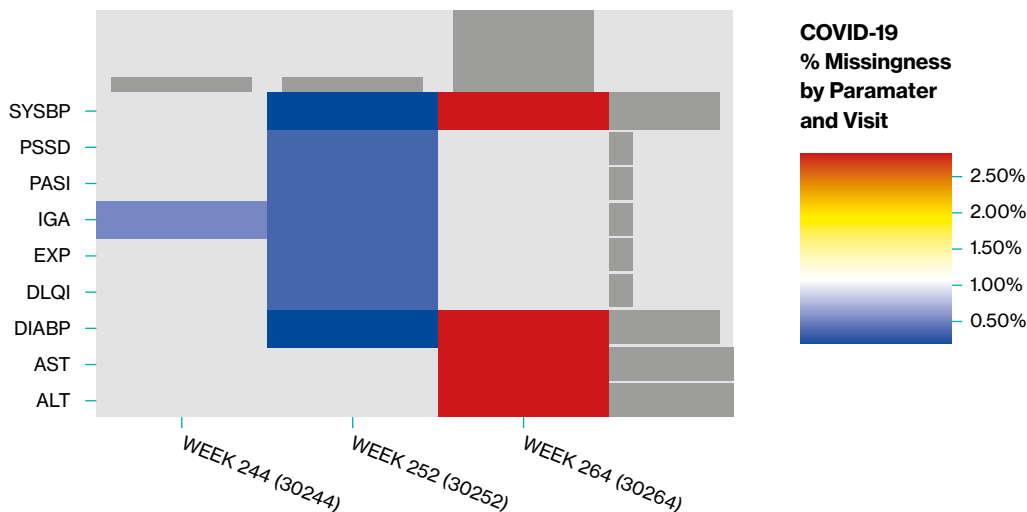
Information collected via Protocol Deviations can be utilised to identify missing planned assessments (primary & secondary efficacy and safety endpoints as examples). A possible way to aid in the quantitative assessment (including visualisation) of missingness is by creating an analysis dataset that contains, per participant and per assessment or assessments of interest, all expected visits/visit dates (actual and missed) through a cut-off point (a predefined point of censoring, i.e. beyond which no assessments are expected. For example, in an oncology study this could be the date of progression. For other TAs, this could be the date of discontinuation for a participant that discontinued study participation). The dataset would have flags or indicator variables to associate missing observation with a reason (e.g. participant infected with AE, temporary measures including site closure). The association of missing data with the reason may need to be derived (based on a well-documented set of data handling rules) unless collected explicitly in the CRF. Such derivation may entail the notion of estimating missing assessment date, creation of visit windows and association with AE onset dates, Protocol Deviation start/end dates, etc. It would also be important to understand how such a dataset may be used not only for ongoing assessment of missingness/ data monitoring activities but also towards reporting (CSR, submission).

Below is an example of visualisations on missing endpoints for a particular study using R-shiny.

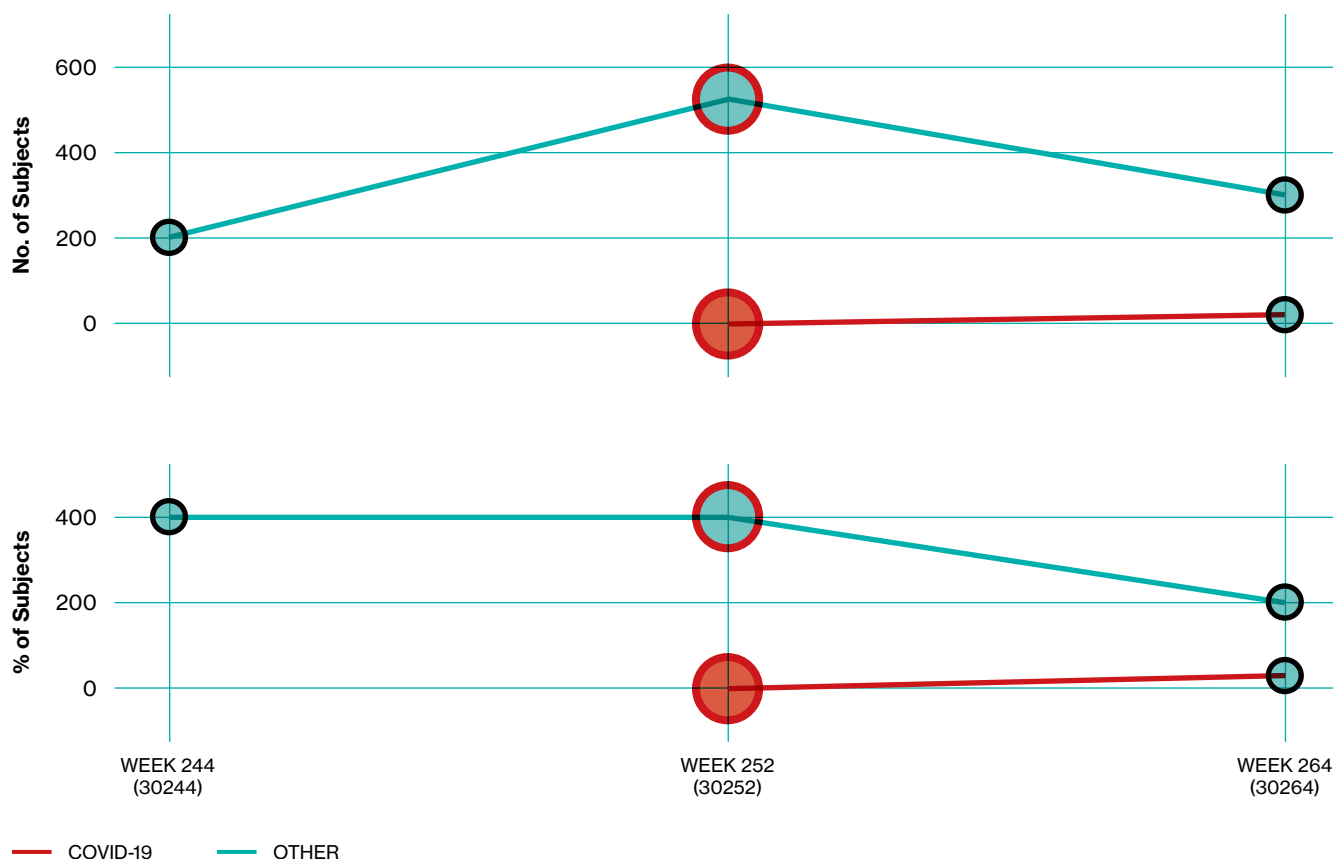
Select the upper bound for % Missing due to COVID-19 by PARAMCD & VISIT



Heatmap of COVID-19 & Missingness by Parameter and Visit



**Missingness Type for the SYSBP Parameter by Visit (No. and % of subjects)†**



Another interesting aspect is to review the lag time between start of event and entry of event into EDC prior to the COVID situation and then during/after the lockdown periods. The results of this can promote discussion with sites to determine individual site status and whether or when the sites move back to a period of normality with regard to event/assessment reporting.

**WILL THE AMOUNT OF MISSING DATA AFFECT POWER?**

If the amount of missing data is thought to affect power, there are several mitigating actions that can be employed. One possible approach is to recruit more patients; however, there may then be a risk of overpowering if the amount of missing data is less than expected. Another option is to amend the protocol to have a longer period of follow-up and/or delay data cuts or database locks. All these options would need to result in a review of the protocol and the need for an amendment. The reporting of event-driven studies may also be affected and require mitigating actions. Further analyses may be required to be added to the statistical analysis plan, either as sensitivity analyses or additional sub-group analyses, which is further discussed in the section below.

**DMCS AND INTERIM ANALYSIS TO IDENTIFY IMPACT ON TRIAL VALIDITY**

To identify the impact of the pandemic on the analysis, it is recommended that an analysis of the trial data is conducted by an independent DMC, which may already exist for the trial. If not,

an independent DMC should preferably be established, following the necessary procedures regarding ethics committees and relevant competent authorities.

- The statistical analysis section of the protocol and/or the statistical analysis plan may need updating based on the findings.
- As a general principle, there are strong scientific reasons to conduct trials as planned and implement changes only when there is a convincing scientific reason that it improves interpretability of results.
- There may be situations, however, where an unplanned (or early) analysis should be considered to minimise the effect of COVID-19 on the interpretation of the data at the risk of having lower power than originally planned. These include situations where:
  - The trial is close to completion.
  - A planned interim analysis is due soon.
  - The trial needs long-term follow-up to observe the primary outcome, especially in cases where enrolment of the trials will be slowing down or even paused during the pandemic.

**ANALYSIS AND REPORTING OF MISSING DATA AT THE CONCLUSION OF A TRIAL**

The FDA and EMA issued guidelines on how the COVID-19 pandemic may affect the conduct of clinical trials. Based on these guidances, here are some principles to consider during analysis and reporting of missing data due to the pandemic.

### 1. Impact on Treatment Effect Evaluation

In order to identify and address concerns about the impact of the pandemic on treatment effect evaluation, sufficient amount of information on pandemic-related measures and whether trial participants or trial conduct were affected, as well as on the subpopulations of exposed/non-exposed, and infected/non-infected participants will be necessary. Some recommendations are:

- Consider whether trial objectives are affected by the pandemic and whether trial estimands need to be modified (e.g. using alternative endpoints)
- Consider design and analysis strategies and ways to handle potentially any altered endpoints, higher measurement variability and missing visits.

### 2. Impact on Validity of the Trials

The external validity of trial outcomes may be affected by the presence of different trial populations: some participants were present in the trial before the start of the pandemic; some during the pandemic while possibly exposed to associated measures; and some after the end of the pandemic.

Consider an analysis of the accumulating trial data in order to evaluate the implications on recruitment, loss of participants during the trial, the ability to record data and to interpret the treatment effect in light of the pre-, during and post-pandemic measures phases.

### 3. Missing Visits/Endpoints

As noted earlier, changes in study visit schedules, missed visits, or participant discontinuations may lead to missing information (e.g. for protocol-specified procedures). Regulatory agencies expect this information to be summarised in the clinical study report by sponsors [2].

A few recommended analyses are included below for consideration. Readers are encouraged to utilise these suggestions as general guiding principles to formulate relevant analyses and displays that best fit their situation.

Any pre-planned analyses that are modified because of the COVID-19 pandemic should be explained, and clarity on assumptions, characterising the type and amount of missing data for each estimator, should be documented. Also, a broad range of sensitivity analyses and methods of estimation of missing data should be considered.

Supplementary analyses using historical clinical trial data or real-world data to support assessment of the impact of missing data elements might be helpful considerations. Summaries of missed dose impacting dosing compliance, and disposition tables showing impact of COVID-19, as well as displays summarising the amount of missing data for primary (and possibly key secondary) efficacy analyses, should be considered. In addition, displays explaining the impact of lab data used as a trial endpoint which cannot be measured in the central lab and has to be done locally will be good considerations. Additional displays exploring further sub-groups (e.g. infected vs non-infected, site vs remote, age categories) should be considered if necessary.

### HOW MUCH MISSING DATA IS TOO MUCH?

The COVID-19 pandemic has created an exceptional situation where study teams may face a large amount of missing data, sometimes critical to study validity and patient safety. The specific reasons and situations were nicely described in previous paragraphs. Strategies proposed by sponsors and sites to address the situation and minimise impact have been many. However, even with best intentions, ensuring 100% data completeness and the necessity to ensure patient, site staff and sponsor staff safety, as well as adherence to country-epidemic regulations, turned out to be difficult or impossible under lockdown.

Regulators supported industry by issuing guidance, and several approaches have been described above on collection, adjusting visit frequency etc., up to proper reporting in CSR. Over time we will see what the final impact has been on study validity.

So how much data missingness is too much?

There is no single answer to that. The answer depends on many factors: tested disease, specific study design, therapy and many others. For example, missed disease evaluation visits may impact oncology study reportable results much more than a vaccine trial. Or missing safety data of a Phase II study testing therapy with a narrow therapeutic window may be more impactful on patient safety oversight compared to a well-established therapy observational study. Each study will need individual evaluation by study teams, and statisticians overseeing statistical plans will have a prominent role in interpretation. The support and collaboration from a clinical data scientist or programmer to operationalise the statistical plans will be critical in this regard.

In addition to methods described in detail in previous chapters, one of the possible techniques to verify missing data impact could be the use of a Quality Tolerance Limits (QTLs)-like approach.

In a nutshell, the QTL concept, defined by ICH GCP E6 R2 [10], has been established to ensure implementation of proactive measures to secure data quality/validity and patient safety at study level.

A QTL is a level, point or value associated with a parameter which is critical to study quality. QTLs are identified at study level and, if exceeded, may indicate systematic issue that can impact subjects' safety or reliability of trial results and as such should trigger an evaluation.

Typically, QTLs are defined at the beginning of the study, are driven by Critical to Quality factors and followed over the course of the study. If there is a risk of crossing the threshold, corrective or preventive action takes place.

When defining QTLs, teams focus on what truly matters for study validity. TransCelerate recommends selecting 3–5 parameters [11]. The threshold setting is driven by statisticians and clinical. The teams use similar studies' historical data and current study statistical plans. If a QTL targets a parameter related to missing data, the teams may immediately assess its impact.

It may be that a sponsor does not currently use QTLs addressing missing data or does not yet have a QTL process in place. Still, a similar approach could be used to set a snapshot of a study and continue oversight of the study moving forward. One-time parameter and threshold identification could be defined using the same methodology and observed moving forward.

Although this approach does not follow the intention of ICH GCP to set QTLs at the beginning of the study, it seems that it can be very practical under the current COVID circumstances. One important thing to remember is that use of QTLs requires formal reporting in CSR. This should not be a problem as sponsors are already required to report on COVID-related issues in CSRs. We recommend TransCelerate guidance on QTLs for those who would like to consider this approach.

## Efficacy

The concept of efficacy can cover a very wide range of endpoints and domains, depending on the therapeutic area and the type of study drug under investigation. It may cover visit-based assessments, e.g. laboratory assessments, RECIST scans, PRO assessments or respiratory assessments, or may also incorporate events, e.g. deaths, relapses, or progression. Impact on statistical inference will be considered, but in-depth analysis of this is not in scope.

We summarise firstly the key points to consider from the FDA and the EMA on the impact of COVID-19 on efficacy in trials which are ongoing during the pandemic.

### FDA GUIDANCE ON COVID-19 IMPACT ON EFFICACY: KEY POINTS [2]

- With respect to efficacy assessments, the FDA recommends consultation with the appropriate FDA review division regarding protocol modifications for the collection of efficacy endpoints, such as use of virtual assessments, delays in assessments, and alternative collection of research-specific specimens, if feasible.
- For individual instances where efficacy endpoints are not collected, the reasons for failing to obtain the efficacy assessment should be documented (e.g. identifying the specific limitation imposed by COVID-19 leading to the inability to perform the protocol-specified assessment).
- If the results of laboratory tests or imaging assessments are the basis for formal hypothesis testing, including primary or secondary efficacy endpoints, sponsors should consult with the relevant FDA review division. For example, disparities in laboratory measurements or imaging protocols will introduce increased variability and thus can affect type I and type II error rates.

In addition, the FDA has recently issued guidance on statistical considerations for proposed changes to trial conduct that may impact the analysis and interpretation of primary or key secondary endpoints in the trial [12]. The FDA recommends that sponsors consult with the relevant FDA review division when considering protocol changes and changes to the statistical analysis plan that may impact the analysis and interpretation of these endpoints.

### EMA GUIDANCE ON COVID-19 IMPACT ON EFFICACY: KEY POINTS [13]

The following is a brief summary of points from the EMA guidance that may impact on efficacy assessments and analysis.

- Impact on recruitment, data collection, analysis, and interpretation of results for each trial will need a thorough case-by-case assessment.
- Risk assessment of the impact of COVID-19 potentially affecting trial participants directly and affecting clinical trial conduct on trial integrity and interpretability is recommended.
- Sponsors are advised to pre-plan how systematic deviations resulting from the measures and individual decisions related to the COVID-19 pandemic are captured.
- Data collection should preferably not stop and should continue as long as possible. However, potential risks for study participants when undergoing study-specific procedures take priority in decisions taken by study participants and health institutes.
- Potential follow-up considerations based on the risk assessment may include the following:
  - proposals to deal with any identified potential sources of bias comprising identification of newly emerging intercurrent events or missing values, or other unforeseeable required changes to trial elements;
  - the need to adjust the trial sample size;
  - recommendations from a trial participant's safety perspective on how to stop, pause or restart the trial;
  - recommendations of additional measures when completing the trial after the pandemic (e.g. validation of outcomes that were measured differently).

### DISCUSSION ON IMPACT OF COVID-19 ON EFFICACY ANALYSES

An estimand is a precise description of the treatment effect reflecting the clinical question posed by a given clinical trial objective [14]. There are five attributes used in the construction of estimands; namely the treatment condition, the population of interest, the patient-level endpoint (variable), the handling of intercurrent events, and the population-level summary measure.

Intercurrent events are events occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest. Examples may include use of rescue medication, discontinuation of treatment, and death.

COVID-19 may introduce additional intercurrent events, and the definition of estimands may need to be revised, or additional estimands may be required to be defined. Therefore, a study may require additional sensitivity analyses. It is very important to perform a full risk assessment for each study and identify a list of efficacy endpoints that may be affected by COVID-19, and therefore may need sensitivity analysis. Subsequently, a list of variables should be identified in order for additional information to be collected on the CRF.

The Pharmaceutical Industry COVID-19 Biostatistics Working Group (Meyer et al., 2020) [15] recommends a forward-looking assessment to anticipate effects which may occur, and that risk assessment should continue and be updated regularly throughout the trial. It is highly recommended to reference this



paper for more detail on risk assessment and estimands. Examples of potential disruptions to efficacy assessments due to COVID-19 which may be defined as intercurrent events could include:

- missed visits/assessments
- remote/virtual visits
- alternative methods of assessment (e.g. local assessments instead of central, RECIST scans being done at different locations)
- missed study drug dosing
- changes in study drug dosing
- diagnosis of COVID-19
- death.

Handling of intercurrent events during the COVID-19 pandemic may need to be adjusted due to a different expected frequency of the event. For example, interruptions to the study drug may have been rarely expected and therefore not previously have had a significant impact on efficacy. However, during the pandemic there have been study-/region-wide drug supply problems and therefore interruptions to the study drug may be more frequent and have a much greater impact.

It is important to ensure that all variables related to these events are collected on the CRF and mapped appropriately into SDTM/ADaM to enable sensitivity analysis to take place. For example, the visit-based information for remote/alternative assessments should be collected in the same visit structure as the visit-based efficacy endpoint in order for it to be merged to the corresponding analysis dataset and flagged appropriately.

There may also be problems in the ability to test the existing primary endpoint, for example hospitals may be unable to conduct certain assessments such as progression-free survival scans or FEV1 assessments. The sample size or the primary estimand may need to be adjusted. Sponsors should ensure enough data is collected to answer the original scientific question posed by the trial.

Sensitivity analysis may need to be conducted by region, by time period (e.g. pre-pandemic vs during pandemic) or by assessment method as COVID-19 may have a high impact on results, e.g. quality of life assessments. Effects such as these may introduce confounding in establishing the efficacy of a drug. In this case it is important to clearly define the regions and the time periods, e.g. what date is considered as the cut-off for the start of the pandemic? This cut-off may differ between country and region. Virtual or remote assessments may cause additional variability in data, and it is recommended [15] that if the conduct/method of an assessment has not already been tested that a pilot testing of a small number of patients occurs before applying this to the entire trial population.

It is recommended that the target population for the efficacy analysis remains as originally planned [15]. However, sensitivity analysis may be required on additional populations, e.g. pre- and post-pandemic randomised patients, or subsets of the full analysis set. As mentioned previously, care should be taken over these definitions to be very specific about any cut-offs used and these should be clearly defined in the protocol and/or SAP.

For studies where a significant number of deaths would have been expected pre-pandemic, or where assessment of death

is required for key endpoints (e.g. overall survival), sponsors need to ensure the reason for death is very accurately captured on the CRF so that sensitivity analysis can be conducted differentiating expected deaths and additional deaths. There needs to be detailed discussions on clear definitions of whether a death is associated with COVID-19 and the specific reasons available for death in the creation of the CRF fields and CRF completion guidelines. If a patient has a positive diagnosis of COVID-19, the death may be solely due to COVID-19, or due to existing medical conditions exacerbated by COVID-19. The patient may also have suspected COVID-19, but there was no ability to test for a confirmed diagnosis. The patient may have died due to problems in an overwhelmed health service but not actually had COVID-19 themselves.

### SDTM GUIDANCE

Missed visits/ assessments	Information on missed visits can be mapped into the custom Visit Events (VE) domain [6]. The VE domain is recommended as SV can only be used to map visits that actually occurred. VE can map missed visits and reasons for the missed visit (FDA recommended [1]). The full list of visit names and numbers should be mapped to VE.VISIT and VE.VISITNUM. The numbers and naming should be consistent with the SV domain such that VISIT and VISITNUM values are the same. VE.VEOCCUR should be marked Y/N if the visit occurred or not, and if not, then the reason should be documented in VE.VEREASOC.  Regarding missing assessments, CDISC [6] recommends the following: If a visit was missed completely, there is no need to represent separately the fact that some assessments scheduled for that visit were missed. If some of the data scheduled for a visit was collected, then the assessments expected but missed at that visit generally should be represented in the relevant domain in with --STAT = "NOT DONE" and the reason for being missing in --REASND.
Remote/ virtual visits and alternative assessments	Information on remote/alternative visits can be mapped into the custom Visit Events (VE) domain [6] (similarly to missed visits), with the mode of contact being specified in the non-standard variable (NSV) VECNTMOD.  If specific assessments require a different type of assessment, it is suggested to record this in the relevant domain in the standard variable --METHOD.
Interruption or discontinuation to study treatment	Last dosing record could represent whether the stoppage of study treatment is permanently discontinued or temporary by populating either EXRSDISC (for permanent discontinuation) or NSV EXRSINT (for temporary discontinuation) [6]. In case of permanent discontinuation, then DSTERM could indicate if discontinuation is COVID-19 related or not using NSV DSEPRELI.  CDISC guidance indicates that temporary interruptions or adjustments to study treatment can be captured by introducing collection of the following NSV variables:  EXRSINT EXEPADJI EXEPINTI EXEPDSCI [6].

Infection/ Diagnosis of COVID-19	<p>If information about COVID-19 infection/symptoms is collected in studies that are ongoing, then CDISC recommends the information be mapped into AE (using the NSV AEEPRLI = Y) [6].</p> <p>Diagnosis date, i.e. specific identification of virus from a collected sample will be mapped into the MB domain.</p> <p>Note: historical infections/diagnosis of COVID-19 should be mapped into MH.</p>
Deaths	<p>Sponsors need to ensure the reason for death is accurately captured on the CRF. If death is part of a key endpoint of the trial (e.g. overall survival), then reasons and dates of death should be mapped to CE as standard. If deaths are not an endpoint of interest, then the information should be mapped to the DS, AE and DD (Death Domain) as standard.</p>

## ADAM GUIDANCE

ADaM datasets may need to be more complex than originally planned due to the additional information required for analysis. This may be present in the form of additional subject-level population flags, additional record-level flags, or additional derived parameters.

Additional population flags can be added to ADSL, in case sensitivity analysis needs to be performed. CDISC guidance already allows for bespoke population flags in the form of ----FL, using Y/N values for each patient. These ADSL flags can then be merged onto each ADaM.

Derived parameters may need to be added for efficacy, e.g. in ADTTE there may be additional parameters required where we censor at date of diagnosis of COVID-19.

The visit windowing in AVISIT/AVISITN should remain the same as planned, but analysis flags can be added to the datasets to give the opportunity to perform sensitivity analysis.

Record-level flags for inclusion in analysis can also be added in each ADaM in the form of CDISC-specified population record-level flags (e.g. FASRFL or PPROTRFL) or bespoke population record-level flags (----RFL). If record-level flags are not specific to a population definition, general flags such as ANL01FL can be added.

## TABLES, FIGURES AND LISTINGS

The key elements leading to TFL production related to trials ongoing during the pandemic are:

- risk assessment
- implementing any required protocol or SAP amendments, in particular adjusting any estimands
- collecting of any additional information on the CRF in an accurate and clearly defined way
- mapping additional information accurately to SDTM
- using the SDTM information to derive additional populations, parameters or record-level flags in ADaM.

The additional information in ADaM can then be used for

selection into the relevant efficacy TFLs including any sensitivity analysis.

## Adverse Events

### DATA COLLECTION

All COVID-19 infections and positive tests should be collected as adverse events in a study. It is also important to note that when collecting adverse events that there has been, in many jurisdictions, limited access to testing. This may make the identification of confirmed cases of COVID-19 infection challenging. Symptoms that could appear to be COVID-19 infection related do not always follow with a positive test. Sponsors will need to exercise caution to ensure that all COVID-19 related adverse events are correctly identified.

On 19 April, MedDRA MSSO, with the approval of the MedDRA Management Committee, released an updated version of MedDRA 23.0 with new COVID-19 terms and revisions included to allow organisations to capture, share and analyse scientific and medical information appropriately. Sponsors were urged to implement this re-released version by 4 May to facilitate this [16]. Consequently, clinical data scientists should ensure, as per normal practice, that all adverse events are coded to the latest MedDRA version. The new list of MedDRA terms contains values for positive and negative COVID-19 tests and for asymptomatic COVID-19. Where tests are performed, it is important for sponsors to consider where this data is captured.

Where sponsors do not recode to the latest version of MedDRA, it will be necessary to identify COVID-19 related adverse events. This may be requiring fixed text to be included in the verbatim, for example a prefix of COVID-19 or by updating the data collection to add an indicator that the event is COVID-19 related. Where sponsors do choose to update the CRF, CDISC has provided guidance on where to include this data in SDTM as part of the Guidance of Ongoing Studies Disrupted by COVID-19 Pandemic [6], in this case the non-standard variable AEEPRLI (Epi/Pandemic Related Indicator) has been proposed.

Given that the list of COVID-related terms added to MedDRA v23.0 [17] provides a search criteria for COVID-19, many sponsors will choose not to update the data collection for adverse events to flag COVID-19 related adverse events as these terms can be identified as part of data analysis activities using customised queries. By using search criteria to identify COVID-19 related terms, sponsors may avoid the need to update implementation of EDC systems for studies, which is particularly relevant where multiple concurrent studies are being run. An additional benefit of this approach is that search criteria will likely evolve as scientific understanding of the virus develops.

### DATA ANALYSIS

The basis for adverse event analysis is described in the PHUSE White Paper “Analysis and Displays Associated with Adverse Events: Focus on Adverse Events in Phase 2-4 Clinical Trials and Integrated Summary Documents” [18]. As noted in Nilsson et al. [8], if the impact of COVID-19 affects treatment arms equally, analyses of adverse events from controlled data can largely remain unchanged. Where an imbalance occurs, different analytical approaches may be needed.

For serious adverse events of COVID-19, these should be reported as such and narratives should be included in the CSR as per usual practice. Sponsors' existing processes will generally be sufficient to gather additional details (e.g. treatments, procedures and tests) relating to the adverse event.

It is important to understand the prevalence of COVID-19 related adverse events in a clinical study and to investigate whether there is an imbalance in the distribution of these events between treatment arms. To enable this, study teams will need to review data for COVID-19 related adverse events. Depending on the number of events, this may be achievable by reviewing existing tables and listings for adverse events, serious adverse events and deaths or by generating new outputs for COVID-19 related adverse events.

COVID-19 related adverse events are highlighted in the list of COVID-19 updates to MedDRA v23.0 developed by MSSO [17]. Some sponsors may choose to create a customised query for these terms based on this list to simplify the identification of such events. To support COVID-19 analyses in the future, it would be beneficial for a MedDRA Standardised MedDRA Query (SMQ) to be developed to consistently identify these events across sponsors. Alternatively, some sponsors will have updated their data collection to allow investigators to flag COVID-19 related adverse events; in this case, additional tables, listings and figures will be required to report this subset of events. Where verbatim terms have not been coded to MedDRA v23.0 some sponsors will have provided instructions on how to report COVID-19 related adverse events by use of key phrases in the event description. This would require manual review and/or programming logic across free text fields, which will create additional challenges due to inconsistencies in data entry.

Depending on the number of cases, the investigation into potential treatment arm imbalances could be achieved with interactive data review, listings or, if sufficient numbers, summary tables. As not all search terms are explicitly related to COVID-19 (e.g. "Quarantine" or "Patient Isolation"), some sponsors may choose to only include events that occur after the known start of the pandemic in their review or customised query. When submitting analyses based on customised queries it is critical that the list of terms used is submitted to the health authorities. Where data is submitted, it is beneficial to submit this as a data file and to document this in the Analysis Data Reviewer's Guide (ADRG) [19] to support the reproducibility of analysis.

Where imbalances in the number of COVID-19 related adverse events between treatments are suspected or identified between treatment arms, it may be necessary to present additional tables, figures and listings for adverse events. These could include extra analyses of deaths, study discontinuations and adjustments to study treatment (e.g. discontinuation or delays to administration) due to COVID-19 adverse events. Per the FDA guidance [2], sponsors may also need to consider further subgroup analyses to assess comorbidities.

Some sponsors may pre-plan analyses on the treatment effect of COVID-19 related events if there is the risk that the mode of action of the investigation product(s) may lead to an imbalance of COVID-19 related events between treatment arms. In these instances, it may be necessary to consider instances of COVID-19 as a safety topic of interest. Detailed considerations on this topic are available in the PHUSE White Paper "Analysis

and Displays Associated with Safety Topics of Interest: Focus on Phase 2-4 Clinical Trials and Integrated Summary Documents" [20]. They may also choose to conduct further analyses on events that occur concurrently with a COVID-19 adverse event. This would allow sponsors to identify any additional signs and symptoms of COVID-19 reported separately from the COVID-19 adverse events based on scientific understanding of the virus and the symptoms caused at the time of the collection of the adverse event.

While we have focussed on the identification of imbalances in the distribution of COVID-19 related adverse events between treatment arms, it may also be necessary to evaluate the impact of COVID-19 on some safety topics of interest. In these instances, per Nilsson et al. [8], it may also be necessary to implement more complex statistical methodologies to evaluate this. It will be necessary to take into account both the number of COVID-19 related adverse events and the timing of these events in any analysis.

## Laboratory Tests

### DATA COLLECTION

Due to the temporary closure of many clinical trial sites and/or restriction of activities in many hospitals, ongoing trial visits may be conducted in ways other than specified per protocol. Visits may be conducted as televisits via phone or video rather than as physical in-person visits. While this is beneficial for checking on subject status and to document changes in adverse events and concomitant medications, televisits do not easily support the collection of laboratory data. If warranted and feasible, the investigator may instruct the subject to visit a local laboratory not affiliated with the trial to undergo sample collection. The investigator should specify for the subject the list of minimum laboratory parameters and any additional parameters of special interest to be reported.

The sponsor company must then determine how to accommodate the data from these additional sources. It may be helpful to consider how often the collection of lab data from non-trial associated laboratories is occurring and the volume of data collected this way in order to determine how best to handle this data.

Scenario Description	Points to Consider
<p><b>Scenario A:</b> Do not include the individual results from the non-trial associated laboratory in the trial database. Use this data to support safety monitoring only. Report any unusual findings as adverse events.</p>	<p>One risk with this scenario is the related individual laboratory results not being stored in the trial database in such a way that they can be included in any analyses or subject profile.</p>
<p><b>Scenario B:</b> Include the individual results from the non-trial associated laboratory in the trial database.</p>	<ul style="list-style-type: none"> <li>• In addition to laboratory results, the laboratory name, sample type, normal ranges, method, fasting status, LOINC code, and other supporting data must also be entered into the trial database in addition to sample collection and results reporting. These data are necessary to ensure proper conversion to standard units, if required.</li> <li>• Consider how the trial database is set up to record this additional data. Some trial databases may already be able to accommodate the reporting of local laboratory results. Other trial databases may not be set up to accommodate local lab results, and a database update would be required. Consideration should be given to the amount of effort required to perform the database update.</li> <li>• Consider how additional parameters not specified by protocol, but reported, would be handled. Will these be included in the trial database?</li> <li>• Consider if and how the data will be used in analyses.</li> </ul>

While the investigator may request the subject visit a local laboratory not affiliated with the trial to undergo sample collection, the sponsor may also request that all subjects undergo such testing. Before recommending this, the sponsor should consider:

- The additional burden **on the subject** to undergo a sample collection at an unplanned location
- The additional burden on the **trial site** to enter this data (normal ranges, methods, LOINC codes, sample collection information, results, etc.)
- The additional burden on the **sponsor company** to perform any database updates and additional data reconciliation and data cleaning activities related to the collection and use of this data, as outlined in scenario B above.

Alternatively, if the site or sponsor misses the opportunity to collect such data, it may result in additional burden **on the sponsor company** and **regulatory reviewer**. If lab summaries are somehow incomplete, it may be more difficult to interpret analyses.

It may be useful to consider a hybrid approach where there is a conscious decision by the medical team at the sponsor company regarding which visits can be missed and for which visits and/or specific laboratory parameters an attempt should be made to collect data via a local laboratory not affiliated with the trial.

If trial sites remain open, it may also be the case that subject visits may occur as scheduled, and sample collection can take place as expected but the samples do not arrive in time at a central laboratory due to shipping delays. In these cases, the sample collection should be documented with reason not done, as per standard procedure.

If any enrolled subjects in the trial are diagnosed as COVID-19 positive, this must be reported as an adverse event and coded with the specific MedDRA defined code. Any additional laboratory data, including COVID-19 test results and/or COVID-19 antibody test results, should be handled as determined within the trial following scenario A or scenario B.

Any additional approaches to collection and reporting of laboratory data which deviate from the original protocol should be documented in the study data management plan.

## DATA MONITORING

Monitoring efforts should focus on the following key areas:

Scenario Description	Points to Consider
Missed visits or visits conducted via televisit, not allowing for laboratory sample collection.	This scenario should likely be documented as a protocol deviation.
<b>Scenario A</b> – Alternative methods used to collect laboratory data and data IS NOT reported in trial database.	Ensure copies of laboratory results are available in the subject’s source documentation. Ensure consistency between any adverse event entries and supporting laboratory result documentation.
<b>Scenario B</b> – Alternative methods used to collect laboratory data and data IS reported in trial database.	Ensure all required data are collected. This may include: <ul style="list-style-type: none"> <li>• Items to support conversion to standard units (sample type, normal ranges, method, age, gender, LOINC code, etc.)</li> <li>• Sample collection details (date and time, laboratory, fasting status, etc.)</li> <li>• All information required to support data reconciliation and data cleaning</li> <li>• Reason certain parameters not reported (if incomplete results reported).</li> </ul>
Samples collected out of window.	This scenario should likely be documented as a protocol deviation.

Updates to the study monitoring plan should be made to outline any additional approaches to collection and reporting of laboratory data which deviate from the original protocol.

## DATA TRANSFORMATION

Data transformation rules and/or programs from source data to SDTM and/or ADaM may require revision to accommodate any data collected and reported under **Scenario B**.

During mapping of source data to SDTM, conversion to standard units may be required to ensure consistency. Additional rules for identification of protocol deviations may be required.

During creation of ADaM data, data handling rules for missing data or data collected out of the planned visit windows should be considered. Will data imputation be performed? If yes, is there any impact on any defined data imputation rules? Will visit windows be defined (if not already specified) or widened (if already specified)? How will partial results be handled (some parameters reported, other parameters missed at the same timepoint)?

Updates to the data transformation specifications, Define-XML documentation, study and/or analysis reviewer’s guides, the statistical analysis plan, and/or data monitoring plan may be required.

## DATA ANALYSIS

The general recommendation of this paper is to keep with the planned, pre-specified analyses as much as possible. Most sponsors likely have planned analyses identical or similar to those provided in PHUSE (2013) [21] and PHUSE (2015) [22]. Consider the balance across treatment groups with respect to number of missed visits, changed visits, alternative data collection methods, protocol deviations, etc. If the impact from modifications required in the trial conduct is minimal and/or is balanced across treatment groups, the pre-specified analyses can likely proceed as planned. In these cases, it is recommended to not overcomplicate the CSR with additional “clutter” and sensitivity analyses that won’t add value to the understanding of

the benefit/risk of the investigational product.

When **Scenario B** is implemented in a trial, consider the handling of additional laboratory parameters. It is recommended to analyse only those analytes specified in the protocol. If additional analytes or additional timepoints are collected, they can be presented in subject listings and subject profiles as unscheduled assessments, but they do not have to be analysed. When creating summaries of the observed data and/or changes from baseline, the study team will need to decide which laboratory measurements can be combined. For some analytes, directly combining the data may not be appropriate. Alternatively, study teams can choose a different analytical approach that allows for combining laboratory measurements from different laboratories. A frequently used technique is to report the data as a percent above/below the normal limit. Alternatively, a normalisation method [23, 24] can be used to combine local and central labs in the analysis.

Special consideration should be given to any laboratory parameters that are efficacy assessments. Here, a trial-specific solution is likely required after discussion within the project or study team. If the parameters are relatively “standard”, such as fasting glucose for a diabetes trial, it may be that these can be easily collected by a local laboratory not affiliated with the trial or collected by the investigative site yet out of window. In other studies, the efficacy parameters may not be so “standard”, in which case consideration should be made regarding how to deal with missing data.

It is anticipated that a good number of planned visits may be missed completely or performed out of window. Changes to the data handling/data imputation/visit windowing rules should be considered if many samples are being thrown out of the analysis. In general, it is likely better to use available results than lose data which are not conforming to such rules. If many samples are collected out of window, it should be considered to change table presentations from by visit presentations and to change to minimum or maximum values. This approach is described in more detail in PHUSE (2013) [21].

For long-term studies, the reader may consider establishing a



specific date window when the COVID-19 pandemic had the most impact on the trial. For this time period, specific algorithms may be applied for this time window only. The pre-planned algorithms may be applied to any data collected pre or post the defined time window of this COVID-19 pandemic. Analyses may also be performed by comparing data of subjects impacted by COVID-19 versus data of subjects not impacted by COVID-19. This may be most applicable to lab analytes of special interest, and not all safety labs.

It is the general recommendation of this paper not to plan on any specific subgroup analyses for signal detection related to any changes in data collection due to the COVID-19 pandemic. If the observation time between treatments is similar due to any disruptions, then a key aspect for signal detection of safety issues is met and additional analyses should not be required. If there was a severe disruption in collection of data or an imbalance across treatment groups, then the reader may consider whether any additional subgroup analyses are appropriate.

If any enrolled subject is diagnosed as COVID-19 positive, the recommendation of this paper is not to perform any additional special analyses unless determined on a trial-specific level. Any COVID-19 related information should be included as part of the subject profile.

Any changes to the pre-specified analyses, additional analyses, new or modified table displays, etc. should be documented in an amendment to the statistical analysis plan or in a supplemental statistical analysis plan.

## Conclusion

The COVID-19 pandemic has affected practically all aspects of life, work and clinical research specifically. It is highly probable that a certain form of epidemic risk will stay with us for a long time and the impact will continue. What will change over time is preparedness of societies and in our environment preparedness of sites and sponsors' clinical operations. The sites may be better ready procedurally to ensure performing safe patients' study visits. Sponsors are already introducing many changes that will allow an increase of remote oversight and more central analytics review. The importance of distance monitoring through electronic health records will also grow providing technology and regulations allow.

Given the scope and breadth of the guidance documents, as well as the multiple changes required to be considered for any given study, it is essential to collaborate with all relevant team members to ensure all updates are managed appropriately as well as to understand any internal sponsor guidelines. It will also be beneficial to discuss proposed approaches with health authorities and individual review divisions as part of taking decisions for individual scenarios.

In general, it is advisable to follow the pre-planned, pre-specified data collection and analyses for the trial, although special consideration may be made depending on subject population, indication, length of trial, and other factors. It will be important for data analysis to understand the distinction between what epidemic-related vs what is typical for the individual study. Where additional data collection is performed it is critical that

all the data can be mapped into SDTM in a usable way. ADaM datasets may need additional population flags, record-level flags and parameters to be derived. These should all be detailed clearly in documentation so that the appropriate analysis can be conducted to answer the original questions posed by the trial.

## Disclaimer

The opinions expressed in this document are those of the authors and should not be construed to represent the opinions of PHUSE members' respective companies or organisations or the FDA's views or policies. The content in this document should not be interpreted as a data standard and/or information required by regulatory authorities.

## References

- [1] PHUSE Data to Knowledge – <https://www.phuse.eu/d2k>
- [2] FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency: Guidance for Industry, Investigators, and Institutional Review Boards – March 2020, Updated on 2 July 2020 – <https://www.fda.gov/media/136238/download>
- [3] Guidance on the Management of Clinical Trials During the COVID-19 (Coronavirus) Pandemic – Version 3 (28 April 2020) – [https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/guidanceclinicaltrials\\_covid19\\_en.pdf](https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/guidanceclinicaltrials_covid19_en.pdf)
- [4] End-to-End Data Flow Today: Current Processes and Challenges – Chris Price (Roche) & Praveen Garg (AstraZeneca) on behalf of the PHUSE Executive Summit (April 2019) – <https://www.phusewiki.org/docs/Summits/End-to-End%20Data%20Flow%20Today%20Current%20Processes%20and%20Challenges.pdf>
- [5] MHRA: Guidance – Managing clinical trials during Coronavirus (COVID-19) – updated 21 May 2020 – <https://www.gov.uk/guidance/managing-clinical-trials-during-coronavirus-covid-19>
- [6] Guidance for Ongoing Studies Disrupted by the COVID-19 Pandemic – <https://www.cdisc.org/system/files/members/standard/ta/COVID-19/Guidance%2520for%2520Ongoing%2520Studies%2520Disrupted%2520by%2520COVID-19.pdf>
- [7] PHUSE (2018): Analyses and Displays Associated with Demographics, Disposition, and Medications – <https://www.phuse.eu/documents/working-groups/deliverables/analyses-displays-associated-with-demographics-disposition-medications-in-phase-2-4-clinical-trials-version-20-02-mar-18-11808.pdf>
- [8] Clinical Trial Drug Safety Assessment for Studies and Submissions Impacted by COVID-19 – Mary Nilsson, Brenda Crowe, Greg Anglin, Greg Ball, Melvin Munsaka, Seta Shahin, Wei Wang – <http://arxiv.org/abs/2006.05502>
- [9] O’Kelly, M. and Ratitch, B. (2014), Clinical Trials with Missing Data: A Guide for Practitioners, New York: John Wiley & Sons – <https://www.wiley.com/en-gb/Clinical+Trials+with+Missing+Data%3A+A+Guide+for+Practitioners-p-9781118460702>
- [10] ICH E6 (R2) Good clinical practice – <https://www.ema.europa.eu/en/ich-e6-r2-good-clinical-practice#current-version---revision-2-section>
- [11] TransCelerate Biopharma Inc. (2017), Risk-Based Quality Management: Quality Tolerance Limits and Risk Reporting – <http://www.transceleratebiopharmainc.com/wp-content/uploads/2017/09/Risk-Based-Quality-Management.pdf>
- [12] Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency Guidance for Industry – FDA Guidance for Industry – June 2020 – <https://www.fda.gov/media/139145/download>
- [13] Points to Consider on Implications of Coronavirus Disease (COVID-19) on Methodological Aspects of Ongoing Clinical Trials – EMA/CHMP – 26 June 2020 – [https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-implications-coronavirus-disease-covid-19-methodological-aspects-ongoing-clinical\\_en-0.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-implications-coronavirus-disease-covid-19-methodological-aspects-ongoing-clinical_en-0.pdf)
- [14] ICH (2019), Addendum on Estimands and Sensitivity Analysis in Clinical Trials to The Guideline on Statistical Principles for Clinical Trials – [https://database.ich.org/sites/default/files/E9-R1\\_Step4\\_Guideline\\_2019\\_1203.pdf](https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf)
- [15] Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic (2020): R. Daniel Meyer, Bohdana Ratitch, Marcel Wolbers, Olga Marchenko, Hui Quan, Daniel Li, Chrissie Fletcher, Xin Li, David Wright, Yue Shentu, Stefan Englert, Wei Shen, Jyotirmoy Dey, Thomas Liu, Ming Zhou, Norman Bohidar, Peng-Liang Zhao, Michael Hale – <https://www.tandfonline.com/doi/full/10.1080/19466315.2020.1779122>
- [16] MedDRA Management Committee Announcement about COVID-19 terms in MedDRA (16 April 2020) – [https://admin.new.meddra.org/sites/default/files/page/documents\\_insert/communication\\_meddrav23-0\\_2020\\_0416\\_0.pdf](https://admin.new.meddra.org/sites/default/files/page/documents_insert/communication_meddrav23-0_2020_0416_0.pdf)
- [17] COVID-19 Related Terms MedDRA 23.0 Update Spreadsheet – <https://www.meddra.org/COVID-19-Related-Terms-MedDRA-23-0-UPDATE-Spreadsheet>
- [18] Analysis and Displays Associated with Adverse Events: Focus on Adverse Events in Phase 2-4 Clinical Trials and Integrated Summary Documents – <https://www.phuse.eu/documents/working-groups/deliverables/adverse-events-white-paper-version-10-03-feb-17-11796-19835.pdf>
- [19] Analysis Data Reviewer’s Guide (ADRG) – <https://www.phuse.eu/css-deliverables>

[20] Analysis and Displays Associated with Safety Topics of Interest: Focus on Phase 2-4 Clinical Trials and Integrated Summary Documents – <https://www.phuse.eu/white-papers>

[21] PHUSE (2013): Analysis and Displays Associated with Measures of Central Tendency – Focus on Vital Sign, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Submission Documents – <https://www.phusewiki.org/docs/Deliverables/%2321657%20PHUSE%20Deliverables%20Aug%202019%20%e2%80%93%20Analyses%20and%20Displays%20Associated%20with%20Measures%20of%20Central%20Tendency.pdf>

[22] PHUSE (2015): Analysis and Displays Associated with Outliers or Shifts from Normal to Abnormal: Focus on Vital Sign, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Submission Documents – <https://www.phuse.eu/documents/working-groups/deliverables/analyses-displays-associated-with-outliers-or-shifts-from-normal-to-abnormal-version-10-10-sept-15-11815.pdf>

[23] From Local Laboratory to Standardisation and beyond: Applying a common grading system. Angelo Tinazzi, Irene Corradino, Enrica Paschetto, Sonia Colombini. SENDO-Tech S.r.l., Milan, Italy – <https://www.phuse.eu/documents/archive-documents/from-local-lab-to-standards-and-beyond-1546.pdf>

[24] Karvanen, J. (2003). The Statistical Basis of Laboratory Data Normalization. Drug Information Journal: DIJ/Drug Information Association, 37, 101–107.

## Acknowledgements

PHUSE would like to express our profound gratitude to the task force participants and their respective companies for their involvement in the development of this resource.

Sangeeta Bhattacharya (J&J), Varia Cartledge (Alexion), Adam Czernik (J&J), Karnika Dalal (Bayer), Anbu Damodaran (Covance), Wendy Dobson (PHUSE), Josephine Fong (Genentech), John Franchino (Roche), Amy Gillespie (Merck), Patty Hegarty (Bayer), Becky Hintze (Pfizer), Matt Jones (Veramed), Kevin Kane (PHASTAR), William Lee (Eli Lilly), Vamshi Matta (Covance), Yvonne Moores (GSK), Mary Nilsson (Eli Lilly), Lavanya Peddibhotla (Covance), Chris Price (Roche), Ying Shi (Eli Lilly), Pritesh Solanki (Merck), Mark Tumelty (GSK), Ojesh Upadhyay (GSK), Chris Wells (Roche)