ADDENDUM ON ESTIMANDS AND SENSITIVITY ANALYSIS IN CLINICAL TRIALS

Topics requiring clarification

Prepared by members of the Estimands Implementation Working Group (EIWG)

Co-authors: Christine Fletcher, David Wright, Stefan Englert, Vivian Lanius, Christian Loesch, James Bell, Chun-Hang Tang

Date: 19 Dec 2025

Background

ICH E9(R1) was finalized in November 2019 and introduced a new framework for describing treatment effects of interest, estimands, aligned to study objectives and ensuring analysis methods, including sensitivity analyses, are aligned to the estimands defined. An EFPIA/EFSPI sponsored Estimand Implementation Working Group (EIWG) was set up at the time the addendum to E9 was finalized with the main purpose of supporting the implementation of ICH E9(R1).

Purpose

This document is a collation of topics requiring clarification that have been identified by the EIWG based on 5+ years of experience implementing ICH E9(R1). The topics have been presented in order of importance and suggestions to address current misunderstandings are provided for consideration, which we believe will provide clarity and improve the understanding of key concepts introduced in the addendum.

We are requesting the ICH E9 WG review these topics presented below and consider developing a Q&A document for ICH E9(R1). Other ICH guidance documents have associated Q&A documents, e.g. ICH E11A, and we feel a Q&A document is now needed for ICH E9(R1) which will further advance how clinical trialists implement the estimand framework.

Topics Requiring Clarification

Topic 1: Clinical Question of Interest

The "Clinical Question of Interest" is not clearly defined in ICH E9(R1), and it is unclear how it differs from the trial objective and primary estimand. Many people (clinicians and statisticians) often understand "Clinical Question of Interest" as a general high-level "purpose of the trial" similar to the study objective. Bell et al. (https://doi.org/10.1002/pst.2129) have proposed detailed clinical objectives and propose that there should be substantial overlap between the description of estimands and concise definitions of objectives. This is often understood that the "Clinical Question of Interest" should be a detailed verbal description of the estimand with all its attributes. Polverejan et al. (https://doi.org/10.1007/s43441-023-00524-2) argue that the "Clinical Question of Interest" is a meaningful and concise definition of the treatment effect, using non-technical language for easy comprehension and that it must consider the clinical context of use. They recommend that it should include the target population, treatment and comparators pertinent to that context and population, and outcome of interest reflecting the qualitative aspect of the treatment effect.

Please provide a definition of "clinical question of interest", and how it is intended to relate to trial objectives and primary estimands.

Topic 2: Definition of intercurrent events

In the glossary, Intercurrent Events are defined as "Events occurring after treatment *initiation* that affect either the interpretation or the *existence of the measurements* associated with the clinical

question of interest". It is unclear whether "after treatment initiation" is necessary or this should be replaced by "after treatment assignment". This is to accommodate that, in some clinical trials, events may occur between treatment assignment and treatment initiation, for example patient withdrawal prior to treatment initiation in a randomized open-label trial or manufacturing delays to administering a treatment after treatment assignment such as in trials investigating CAR-T therapies. There should be an opportunity to consider the potential for intercurrent events that might occur after treatment assignment.

It is also unclear what is meant by "existence of the measurements". The term "existence" needs clarity and further explanation and needs to be more explicit in linking this to an underlying biological/clinical property that leads to a measurement being possible. There is some confusion whether the addendum was only focusing on whether a measurement may not exist because the biological/clinical property can no longer be observed, e.g. blood pressure of a dead patient, or if this is also referring to situations when a scheduled visit was missed so no observations were collected. It seems that the addendum was trying to make a distinction, but this is not very clear in the context of intercurrent events. Events that occur after treatment initiation that only lead to missing data or estimation challenges should be discussed within the analysis considerations and not intercurrent events in the estimand. More clarity is needed to ensure terminal events beyond which the outcome does not exist are considered distinct from failures in study conduct to obtain a measurement.

Some people argue that "Improper handling of laboratory samples" which results in missing laboratory assessments are intercurrent events as they seem to fulfil this definition. Others argue that the outcomes do exist and that they would have been meaningful for the analysis thus this is a missing data issue without an intercurrent event.

Please clarify how to identify intercurrent events where there is a gap between treatment assignment and treatment initiation, and what is meant by 'existence of the measurements'.

Topic 3: Determining which intercurrent events to identify

More guidance would be beneficial regarding the selection of intercurrent events to identify for an estimand. The addendum suggests that the estimand should be clinically relevant. However, it is unclear how this should be applied. For example, the terminal event of death occurring before the variable is measured could be an intercurrent event for many clinical trials. Death seems to be clinically relevant in a broad range of clinical conditions. Yet, death as an event during a clinical trial might not be always expected. Thus, whether to identify death as an intercurrent events does not seem to be based on a clinical argument but rather a pragmatic one based on the expected frequency of death occurring in the clinical setting. There has been no consensus on whether intercurrent events should be selected based on expected frequency or clinical relevance.

Related to this, some feel the treatment policy strategy for handling intercurrent events is different from the other strategies described in the addendum in the sense that the treatment policy strategy can account for intercurrent events that are not explicitly identified in the estimand or may be unknown. This contrasts with the other strategies to handle intercurrent events as the intercurrent events have to be specified explicitly.

There is a lack of clarity if all theoretical intercurrent events that may be discussed initially in discussing estimands need to be identified even if they occur with low or zero frequency. Or those that are most relevant for the estimand(s) of interest should be included in the final estimand(s) of

interest, recognizing there may be a step of consolidating all potential intercurrent events before determining which ones are most relevant to address the clinical question of interest.

A robust multi-disciplinary discussion in the study design stage is needed to discuss what events could occur after treatment assignment or treatment initiation (whichever is most relevant) taking into account the disease setting such that measurements to be taken from participants could be confounded leading to an impact on the interpretation of the variable. Thus, events that occur post treatment assignment or treatment initiation and are not expected to affect the interpretation of the variable and do not prevent measurements of ***the underlying biological/clinical property*** to be observed are not expected to be identified as intercurrent events.

Please clarify how to determine which intercurrent events should be identified for estimand(s) of interest and if death should always be referenced as a potential intercurrent event.

Topic 4: Incorporating intercurrent events in estimand descriptions

In section A3.3 the following paragraph describing the Estimand Attributes is ambiguous in terms of how the intercurrent events and strategies to address them should be described in the estimand definition: "Precise specifications of treatment, population and variable are likely to address many of the intercurrent events considered in sponsor and regulator discussions of the clinical question of interest. The clinical question of interest in respect of any other intercurrent events will usually be reflected using the strategies introduced as treatment policy, hypothetical or while on treatment."

A literal interpretation of the above may be that most critical (but not all) intercurrent events should be mentioned directly as part of the specification of the "treatment", "population", and "variable" attributes, and that remaining ("other") intercurrent events (not already mentioned) should be specified in a separate attribute "other intercurrent events". It is thus unclear if all intercurrent events need to be explicitly stated as intercurrent events or if they are incorporated into an attribute they no longer need to be identified as an intercurrent event.

Please clarify if all intercurrent events need to be explicitly stated or if they are incorporated into an attribute they no longer need to be identified as an intercurrent event.

Topic 5: Supplementary analyses

In section A5.3, supplementary analyses are defined by the ICH E9(R1) Addendum as other analyses that are conducted to more fully investigate and understand the trial data. There is confusion about how supplementary analyses differ from analyses of further secondary or exploratory estimands. There are several interpretations of the concept of supplementary analysis as follows:

- Supplementary *analyses* are those investigating different estimators for same estimand. They are different than sensitivity analyses, which stress-test the main estimator assumptions or limitations of data. For example, the main analysis uses an analysis of covariance and a supplementary analysis uses a repeated measures mixed model.
- Supplementary analyses are analyses supporting a clinical question where at least one attribute of the estimand is changed. For example, the variable (endpoint) has a slightly different definition to that being used for an estimand but everything else is staying the same including using the same methods for estimation.

- Supplementary analyses support the interpretation of a specific estimate for an estimand. For example, providing a comparison between treatment groups of the proportion and timing of patients who use additional or alternative medication, when this use is defined as an intercurrent event.

Please clarify supplementary analyses.

Topic 6: Analysis Sets

ICH E9(R1) covers "*issues related to the concept of analysis sets*". It states that "*[c]larity is introduced by carefully defining the treatment effect of interest in a way that determines both the population of subjects to be included in the estimation of that treatment effect and the observations from each subject to be included in the analysis considering the occurrence of intercurrent events*".

Clearly identifying both (a) which of the observed data from (b) the subjects to be included in an analysis after strategies for addressing intercurrent events have been implemented is important for the estimation of an estimand. However, some people argue that the above statements in ICH E9(R1) imply that the term "analysis set" defined as "*The* set of subjects *whose data are to be included in the main analyses* ..." in ICH E9 (and used as such in several other (regulatory) guidance documents) *has now been replaced* by a broader definition based on the set of subjects and the set of individual observations to be included from each subject.

Prior to ICH E9(R1) the full analysis set (FAS) was typically created by defining which participants would be included in the FAS. The definition of the FAS set would not typically include any specific criteria regarding which of the data observed would be used for the analysis. This approach was formally adopted in protocol templates, e.g. the TransCelerate Common Protocol Template. However, as described in ICH E9 "*The intention-to-treat (see Glossary) principle implies that the primary analysis should include all randomised subjects. Compliance with this principle would necessitate complete follow-up of all randomised subjects for study outcomes*" thus there was an implicit expectation that an analysis aligned to the ITT principle would include all observed data for all randomized participants as identified by the FAS. ICH E9 acknowledged the challenges in achieving this noting "*In practice this ideal may be difficult to achieve, for reasons to be described. In this document the term 'full analysis set' is used to describe the analysis set which is as complete as possible and as close as possible to the intention-to-treat ideal of including all randomised subjects*".

Adopting ICH E9(R1) has led some people continuing to define analysis sets based on which participants will be included (e.g. all participants who are randomised and receive at least one dose) and to subsequently clarify which analysis set will form the basis for estimation for a particular estimand. A second step would then be taken to identify which of the observed data points will be used for estimation. This is often achieved by including additional variables and/or additional records in analysis data sets to identify whether observed values are used, ignored or imputed in the estimation of an estimand aligned to the defined strategies for addressing intercurrent events. This approach enables all observed data to be included in the analysis data set and the set of data to be used for estimation is subsequently identified using selection criteria on the records that are used in the analysis. For other people implementing ICH E9(R1) there is uncertainty in whether regulators intended for sponsors to create separate analysis sets for each estimand and whether this is supposed to contain only the data to be used for estimation after implementing strategies for intercurrent events.

Different interpretations of the term “analysis set” and how to create them are leading to some confusion. Please clarify the intended meaning of the term “analysis set”, including whether it refers only to the set of subjects or also to the set of observations per subject, taking into consideration descriptions in protocols and SAPs, other existing (regulatory) guidance documents, as well as practical implementation considerations (e.g. programming specifications).

Please clarify what is expected when defining analysis sets in protocols and SAPs and provide guidance on practical considerations when creating analysis sets.

Additional topics

In addition to the topics discussed in detail above, we propose seeking further clarification on some other points that recurrently come up in practice, see below.

- **Clarification of the definition of the treatment policy strategy**
  There seems to be a contradiction in the current definition for treatment policy: While it states that *“the occurrence of the intercurrent event is considered irrelevant in defining the treatment effect of interest”,* in the next paragraph it reads that *“the intercurrent event is considered to be part of the treatments being compared”*. Since treatments should be assumed to, at least potentially, be causal, intercurrent events cannot be irrelevant or ignored. In addition, we suggest clarifying the definition of the treatment policy strategy so that it focuses on the meaning of the corresponding clinical question, rather than data usage/estimation.

- **Clarification of competing risks-type intercurrent events**
  Intercurrent events that prevent the existence of outcomes are known in statistical literature (especially for time-to-event outcomes) as ‘competing events’. Competing intercurrent events are conceptually different from those intercurrent events that affect the interpretation of the measurement associated with the variable of interest. There are different consequences for available handling strategies (e.g., treatment policy is not available for competing intercurrent events but is for others). Can these consequences be discussed in more detail in the Q&A? In this context, further clarification on the application of relevant intercurrent event strategies in situations with intercurrent events which are considered competing risks (e.g., whilst on treatment/whilst at risk/whilst alive ... vs hypothetical) are very much appreciated. This is a recurrent topic with strong but often opposing views in teams but also in discussion with regulators. This issue is partially enhanced by the fact that the language used for the handling of (competing) intercurrent events is confusing and/or mis-interpreted. Clarification on the meaning and language for applying a treatment policy, hypothetical, composite, or while on treatment strategy for competing intercurrent events is very much needed.

- **Clarifications around the population-level summary**
  The population-level summary:
  - can summarize the variable separately for each treatment condition (single-arm summaries) and can also cover all types of contrasts for the comparison between treatment conditions, e.g. comparing between 2 treatment conditions, comparing pooled active treatments to a control.

- can be multidimensional or a function, see also Mütze T, Bell J, Englert S, Hougaard P, Jackson D, Lanius V, Ravn H (2025). Principles for Defining Estimands in Clinical Trials – A Proposal. Pharmaceutical Statistics, 24: e2432. https://doi.org/10.1002/pst.2432.
- does not necessarily need to be derived from the same analysis as used for decision making (e.g., that provides the p-value).

Clarification would be helpful in these aspects.

- **Clarification on the role non-statisticians have in the choice of estimand(s)**

As the addendum is linked to ICH E9 and Statistical Principles in Clinical Trials many clinical researchers, both in sponsor companies and in regulatory agencies, view estimands to be statistical in nature. As such it has been challenging having non-statisticians engage in the important conversations required in the choice of estimand(s) aligned to the clinical objectives and clinical questions of interest. Further emphasis is needed on the key roles non-statisticians, in particular, clinicians have in these discussions.

In addition, some sponsors have shared examples where, despite agreements made with regulatory agencies in clinical trial protocols on the choice of primary estimands, regulators have asked for post-hoc estimands during regulatory review of submissions and which have been prioritised over pre-specified estimands in regulatory decision making. ICH E9(R1) was developed to ensure sponsors and regulators are aligned in the choice of estimands at the design stage but unfortunately this is not always achieved. Additional clarification on aligning on estimands during the design stage is needed and what is the role of any post-hoc estimands if proposed during the regulatory review process.

- **Clarification on the acceptance of all strategies described in ICH E9(R1) to address intercurrent events**

The addendum introduces 5 different strategies to address intercurrent events and notes there will still be a preference by regulators for using treatment policy strategies for addressing intercurrent events. The addendum also recognised there may be situations where different strategies for addressing intercurrent events might be needed to address important clinical questions of interest. In addition, some sponsors have shared examples where regulators have asked to use treatment policy strategies to address intercurrent events even if their clinical guidance requires a different strategy, such as composite, to be used. Further clarification would be helpful if all the strategies described in ICH E9(R1) continue to be appropriate for consideration in the choice of estimands where such strategies align to address important clinical questions.