DIA

ANALYTICAL REPORT

Defining Clinical Trial Estimands: A Practical Guide for Study Teams with Examples Based on a Psychiatric Disorder

Elena Polverejan[1] · Michael O'Kelly[2] · Nanco Hefting[3] · Jonathan D. Norton[4] · Pilar Lim[1] · Marc K. Walton[5]

Abstract

While the ICH E9(R1) Addendum on "Estimands and Sensitivity Analysis in Clinical Trials" was released in late 2019, the widespread implementation of defining and reporting estimands across clinical trials is still in progress and the engagement of non-statistical functions in this process is also in progress. Case studies are sought after, especially those with documented clinical and regulatory feedback. This paper describes an interdisciplinary process for implementing the estimand framework, devised by the Estimands and Missing Data Working Group (a group with clinical, statistical, and regulatory representation) of the International Society for CNS Clinical Trials and Methodology. This process is illustrated by specific examples using various types of hypothetical trials evaluating a treatment for major depressive disorder. Each of the estimand examples follows the same template and features all steps of the proposed process, including identifying the trial stakeholder(s), the decisions they need to make about the investigated treatment in their specific role and the questions that would support their decision making. Each of the five strategies for handling intercurrent events are addressed in at least one example; the featured endpoints are also diverse, including continuous, binary and time to event. Several examples are presented that include specifications for a potential trial design, key trial implementation elements needed to address the estimand, and main and sensitivity estimator specifications. Ultimately this paper highlights the need to incorporate multi-disciplinary collaborations into implementing the ICH E9(R1) framework.

Keywords ICH E9(R1) · Treatment effect · Intercurrent events · Missing data · Stakeholder · Estimator · Depression · Major depressive disorder

Introduction

Clinical trials were traditionally planned as follows: a general trial objective was stated, then the trial design, analysis sets, and statistical methods determined how the treatment effect was estimated. This approach was not optimal, because the

✉ Elena Polverejan
  epolvere@its.jnj.com

1 Statistics and Decision Sciences, Janssen Pharmaceuticals - Johnson & Johnson, 1125 Trenton-Harbourton Rd, Titusville, NJ 08560, USA

2 Center for Statistics in Drug Development, IQVIA, Dublin 3, Ireland

3 Global Clinical Development, Therapeutic Area Psychiatry, H. Lundbeck A/S, Valby, Denmark

4 Statistical & Quantitative Sciences, Takeda Pharmaceuticals U.S.A., Inc., Lexington, MA, USA

5 Quantitative Sciences Consulting, Statistics and Decision Sciences, Janssen Pharmaceuticals - Johnson & Johnson, Titusville, NJ, USA

definition of what was being estimated by the trial was either not stated clearly or not stated at all.

The ICH E9(R1) Addendum [1] on "Estimands and Sensitivity Analysis in Clinical Trials", released in 2019, (hereafter referred to as "the Addendum") recommends a change in the process of planning, design, conduct and reporting of clinical trials. The Addendum emphasizes that to properly inform decision-making by various stakeholders and to provide clear descriptions of benefits and risks of a treatment, it is important to have precise descriptions of the treatment effects of interest reflecting clinical questions posed by trial objectives (i.e., the estimands) that are clearly understood and relevant to support the decision(s) to be made by the stakeholders. Estimands must be documented in the protocol; trial design and all aspects of trial conduct and the planned analyses flow from their specification. As pragmatic considerations may impinge on the feasibility of estimating a specified estimand, this process will, in practice, be iterative.

The Estimands and Missing Data Working Group of the International Society for CNS Clinical Trials and

Methodology (ISCTM Estimand WG) includes members representing both clinical and statistical functions, with both trial and regulatory experience. This working group had the objective to develop an interdisciplinary process for implementing the estimand framework in the planning stage of a clinical trial. The current paper describes such a process, illustrated by specific examples using hypothetical trials evaluating a treatment for major depressive disorder (MDD). The description of this process and the examples are intended to be a practical aid to clinical trial teams in applying the recommendations of the Addendum to clinical trials across many disease areas.

Section "Process for Selecting and Constructing Estimands" of this paper describes the recommended process for selecting and constructing estimands and highlights key points regarding the estimand attributes. Section "Process for Selecting an Estimator Aligned with an Estimand" describes the process of selecting an estimator aligned with an estimand. Section "Estimand Examples for Major Depressive Disorder" presents multiple examples of estimands for MDD, some with examples of aligned estimators. Section "Discussion" includes discussion points and further thoughts on this topic.

Process for Selecting and Constructing Estimands

As noted in the Addendum, the purpose of a study is to support decision-making by one or more stakeholders who will use the study results. The precise question(s) each stakeholder needs to answer to support their decision-making can be different, and thus different estimands could be defined for each stakeholder identified for a trial.

The ISCTM Estimand WG recommends the following steps in applying the estimand framework:

- Identify stakeholder(s)
- State decision(s) to be made by each stakeholder
- Define objective(s)
- Under each objective supporting main decision making:

  – Formulate the clinical question of interest:

    • Consider the clinical context
    • Consider potential intercurrent events (ICEs) and how they relate to the question

  – Define the corresponding estimand
  – Justify the utility of the selected question and corresponding estimand to the specific stakeholder(s).

This process may, in practice, be iterative. If an estimand is determined not to be estimable, a relevant alternative question of interest that is aligned with the selected objective should be sought.

Identify Stakeholder(s) and Decision(s) to be Made

There are often a variety of stakeholders who will make decisions based on the results of a clinical trial. Health authority agencies (HAAs, such as FDA, EMA, Health Canada, PMDA etc.) might for example need to decide whether a study contributes substantial evidence of short-term efficacy for a new treatment or that a new treatment is effective as maintenance treatment after an initial short-term response. A company developing a new drug might for example need to determine whether a study provides enough evidence of efficacy to decide on continuing its development. Payers might need to determine whether a study contributes substantial evidence of clinically meaningful patient-level benefit for a new drug or whether the decision to prescribe a new drug is more clinically effective over a long-term period than the decision to prescribe another well-established drug. Eventually payers make decisions on whether to include a drug in a formulary, and what level of payment to provide in relation to available products. Physicians and patients will need enough information to enable their individual decision-making on starting a treatment. This might include answering the questions: what benefit can be expected in patients who could adhere to treatment? How likely is it that the treatment would be adhered to?

Estimand examples in "Estimand Examples for Major Depressive Disorder" section highlight the variety of stakeholders for a study and the decisions they need to make. While these examples highlight decisions on the efficacy of a new treatment, such decisions are complemented in practice by those based on safety and risk–benefit evaluations.

Define an Objective(s)

Each objective should support the stakeholder's decision making. For example, if the decision for a HAA is to determine if the study contributes substantial evidence of efficacy for a new monotherapy drug for MDD, the following objective supports this decision (see Estimand 1 example in "Estimand Examples for Major Depressive Disorder" section): *To assess the superiority of new drug versus placebo in short-term symptom reduction when given as monotherapy treatment in MDD patients*. The statistical hypotheses for an endpoint (e.g., superiority or non-inferiority) or the statistical decision rules (e.g., Go/No Go decision rules) relate to the chosen objectives. A trial objective should mention both the treatment conditions that are being compared and the target population for treatment, both being attributes of an estimand (as discussed in "Define the Estimand" section).

Multiple objectives typically inform each stakeholder's decision making. Protocol templates [2, 3] require that the

included objectives reference all endpoints selected for the trial. These objectives are usually prioritized for the trial as primary, key secondary, other secondary or exploratory to distinguish those used for main decisions (primary and key secondary), and those that have supportive or other roles. This distinction is especially important in the regulatory setting. Of note, it is possible for multiple objectives to reference the same endpoint (e.g., for different target populations).

Formulate the Clinical Question of Interest, Define the Corresponding Estimand, and Justify Their Utility to the Stakeholder

As mentioned above, an objective is a general statement of what supports a stakeholder's decision. The clinical question of interest is a meaningful and concise definition of the treatment effect, best formulated using natural, non-technical language for easy comprehension; it is paired with a formal, detailed definition of the corresponding estimand. They must be relevant to the stakeholder and have their utility justified. All the estimand examples from "Estimand Examples for Major Depressive Disorder" section include these three components.

Formulate the Clinical Question of Interest

The formulation of the clinical question of interest must consider the clinical context of use. This involves consideration of:

- Target population (including typical comorbidities and behaviors)
- Treatment and comparators pertinent to that context and population (including the availability and effectiveness of alternative treatments in the target population)
- Outcome of interest, reflecting the qualitative aspect of the treatment effect (e.g., achieving or avoiding a certain discrete outcome such as treatment success or failure, time to an outcome, change in a continuous score) as well as its temporal aspect (e.g., effect at a fixed time point, over a fixed period, at a variable point in time, over a variable period).

When these have been carefully specified, potential intercurrent events (ICEs) can be considered. ICEs [1] 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 (e.g. treatment discontinuation, starting alternative treatments, death; see Sect. Identify ICEs). Once the ICEs pertinent to the clinical context are identified, a study team can formulate a precise clinical question of interest, for example *"For a patient with MDD, what would be the expected effect of prescribing drug X on depression severity at Week 8, were no other antidepressant medications available?"* While this target treatment effect will be formalized in the estimand definition, formulating the clinical question of interest is an important step as it allows a cross-disciplinary discussion in the study team.

The clinical question of interest formulation needs to capture a clear, specific treatment effect of interest relative to each group of identified ICEs. When the estimand is defined (see Sect. "Define the Estimand"), estimand attributes including the strategies selected for the identified ICEs (see Sect. ICE-Handling Strategies, Table 1) will be linked to the clinical question of interest. Examples of types of clinical question of interest formulations (implying different ICE strategies) are presented below:

- Treatment effect *under the assignment to either experimental treatment or placebo, regardless of ICE*—**Treatment policy** strategy
- Treatment effect under a counterfactual scenario (e.g., *as if patients would continue treatment as assigned* or *as if patients would not start other pharmacological treatments for MDD as they were not available*)—**Hypothetical** strategy
- Treatment effect on the *likelihood of a patient experiencing a treatment response*, where the response definition incorporates the ICE (e.g., patient with ICE is considered as non-responder)—**Composite Variable** strategy
- Treatment effect *while treatment is being taken*—**While on treatment** strategy
- Treatment effect in a stratum of patients who would/would not experience the ICE (e.g., in *MDD patients who would adhere to drug X as prescribed for Y weeks*)—**Principal Stratum** strategy.

The examples above are not exhaustive; other language and formulations that link to different ICE strategies could also be used in the question of interest.

The question should be formulated concisely as possible to serve as a guide for the specification of the estimand. Therefore, when formulating the clinical question of interest, some attributes of the corresponding estimand need not be detailed (e.g., exact endpoint, such as the method/scale of capturing depression severity, or exact population-level summary) or may be implied by the description of the effect (e.g., "expected effect" may imply that the population-level summary will be a difference of means).

Define the Estimand

The estimand is a formal, operationalized expression of the clinical question of interest, constructed with the following attributes (see Section A.3.3 of the Addendum):

- **Treatment condition of interest and Alternative treatment condition** The interventions being compared. Here, not only the experimental treatment (versus control, if applicable) should be specified but the planned treatment regimen as a whole, including (if applicable) the recommended use of additional or background treatment and/or the strategies for handling ICEs related to the treatment regimen.
- **Population** The population targeted by the clinical question of interest. (It can also reflect a population defined by membership in a principal stratum—see Table 1 for definition of the Principal Stratum strategy). This differs from the analysis set (e.g., all randomized participants), referred to in the past as the analysis population, which should be described under the estimator specifications.
- **Variable (or endpoint)** A value that can be measured in individual patients that is required to address the clinical question, e.g., change from baseline to time X in a measure, time to an event, a binary responder variable. It cannot be a proportion, for example, as this cannot be measured per patient. It can take into account ICEs if the Composite Variable strategy is used, or it can reflect the patient-dependent treatment duration if the While on Treatment strategy is used.
- **Population-level summary** The population-level quantity (derived from the patient-level Variable) that provides a basis for comparisons between treatment conditions and quantifies the treatment effect.
- **ICEs and corresponding strategies** Here, strictly speaking, only the ICEs not covered in the other attributes should be specified together with the strategies used to handle them. However, to improve clarity in this implementation phase, we prefer to list all ICEs and corresponding strategies, including those reflected in other estimand attributes. Patients could experience overlapping ICEs and, if these ICEs are addressed with different strategies, the priority order of applying these strategies must be specified. This will depend on the clinical context; for example, the composite variable strategy will most likely have a higher priority over strategies such as treatment policy or hypothetical (see Sect. ICE-Handling Strategies).

The Addendum recommends at a minimum that estimands for all trial objectives that are likely to support regulatory decisions (such as those related to primary and key secondary endpoints) be defined and specified explicitly. If the trial is to serve multiple stakeholders with different questions of interest, estimands for each stakeholder should be formulated in the protocol or in other prospectively written associated documents. A particular estimand might be of interest to multiple stakeholders, as reflected in some of the estimand examples from "Estimand Examples for Major Depressive Disorder" section.

The following sub-sections provide additional details on the identification of ICEs and on the types of available strategies for addressing ICEs.

Identify ICEs

All foreseeable ICEs that are likely to be relevant for a trial are to be identified when planning the trial (see Section A.3.1. of the Addendum). The applicable ICEs depend on the specific setting of the trial, but the following is a list of ICEs that are often encountered based on authors' experience:

- ICEs related to the study treatment:
  - Treatment discontinuation (Tx DC)
  - Change in planned dosage or frequency of administration
  - Treatment non-adherence (i.e., intermittent or partial adherence)
- ICEs related to initiation, adjustment or discontinuation of treatments that are concomitantly taken with the study treatment and may influence the outcome of interest
- Changes in how the outcome of interest is measured (e.g., use of uncertified rater or scale, switching to remote assessment)
- ICEs precluding the existence of values after the event, such as death.

Events could also occur that impact the validity or interpretability of the outcome measurement tool. For example, a cerebrovascular accident could reduce the reliability of assessment of psychomotor impairments attributable to a major depressive episode.

Disease specific regulatory guidance documents for Industry have started to recommend ICEs of interest and strategies to address them, such as the FDA guidance [4] for Chronic Rhinosinusitis with Nasal Polyps or the EMA Guideline [5] on the clinical investigation of medicines for the treatment of Alzheimer's disease.

On rare occasions a major unforeseen source of ICEs may occur. For example, at the time of writing, clinical trials are being impacted by the COVID-19 pandemic and by the war in Ukraine, resulting in disruption to the provision of drugs, changes to methods of assessment, but also affecting the health of the study subjects, and leading to changes in circumstances (individual or societal) affecting the relationship between disease severity and impairment of function or the reliability or validity of measures designed for use under normal social conditions. In these situations, protocols and other study documents such as Statistical Analysis Plans (SAPs) must be amended to address these unforeseen, major, broadly occurring ICEs [6–9].

Each type of ICE could be considered as a unified event or could be further divided into sub-categories. For example, Tx DC due to different reasons (e.g., due to adverse events, lack of efficacy, or other reasons, such as site closures or other administrative reasons) could be considered as one or as different ICEs depending on reason for Tx DC; likewise different severities of the same event such as low/moderate versus severe treatment non-adherence could be considered separately. Different strategies could then be used if these different events are addressed differently in the clinical question of interest.

ICEs are not synonymous with missing data. Indeed, it is usually desirable to collect data after ICEs, and there are data that are missing without (known) occurrence of ICEs. Study withdrawal is not considered by the Addendum as an ICE. Rather, it is a study event leading to *missing data* (i.e., data that would be meaningful for the analysis of a given estimand but were not collected). Some ICEs might be immediately followed by missing data (which could also be intermittent), while others not. The ICE of death cannot lead to missing data as no measurements exist and can be collected after death.

ICE-Handling Strategies

ICEs can be addressed by several potential strategies that are described in Section A.3.2. of the Addendum. Table 1 describes each of the five strategies, points to consider on the use of each strategy, and additional considerations on estimation (see Sect. Process for Selecting an Estimator Aligned with an Estimand on the process for selecting an estimator aligned with an estimand). The formulation of the clinical question of interest should drive the selection of strategies addressing the identified ICEs. This requires a collaborative effort across disciplines and is not an exercise for statisticians only.

Process for Selecting an Estimator Aligned with an Estimand

For each of the estimands, an aligned method of analysis, or estimator [1], should be implemented that is able to provide an estimate on which reliable interpretation can be based.

Once an estimand is defined and the aligned estimator is selected with the chosen assumptions, the following elements are recommended to be included in the estimator specification:

- Define the estimand and estimator aligned analysis set, specifying not only what trial participants are included (e.g., all randomized) but the selection of measurements to be used for each participant.

    Here, specify what data are not used or missing or sometimes not existing, including:

- Data not used—Data that may be collected but are not used for the estimator chosen for this estimand, for example the endpoint values collected after an ICE and replaced by imputation;
- Missing data—Data that would have been useful but could not be collected (e.g., due to withdrawal from the study or intermittent missing)—considered the "true" missing data by the Addendum;
- Data not existing—such as data after death or, for Principal Stratum estimators, data on the occurrence of ICEs had the patient been assigned to other treatment instead.

- Specify the main estimator for this estimand, including:
    - Assumptions for data not used and missing data; these assumptions, whether the data is treated as missing due to an ICE or simply missing because not collected, inform the scenarios analyzed by the statistical model, and may for example lead to censoring, imputation or generation of a composite outcome.
    - Statistical model and its assumptions (e.g. proportional hazard assumption for Cox regression).

- Specify the sensitivity estimator(s) for this estimand, ensuring that the same estimand is targeted and stating how elements and assumptions differ from those of the main estimator.

Extensive details on selecting estimators aligned with an estimand are provided in Mallinckrodt et al. [20]. Of note, as this is a rapidly evolving field, it is likely that any recommendations beyond those of principle could be superseded. Mitroiu et al. [21] provided a summary of what analysis methods have been commonly used in short-term depression studies, mapping estimands to these methods.

The main estimator produces an estimate for the estimand population-level summary, a clinically understandable estimate of the amount of clinical benefit (or risk, for a safety variable) that was associated with the treatment. This is often loosely referred to as the 'study result'. As mentioned in Section "Define an Objective(s)", an objective often includes the statistical hypotheses for an endpoint (e.g., superiority or non-inferiority) or the statistical decision rules. Ideally, the analysis used for decision making should be same as the main estimator or at least with similar assumptions. However, it is possible for the analysis used for decision making to be different than the main estimator, especially for the binary and time to event endpoints. As an example, the population-level summary of hazard ratio for a time to event endpoint can estimate the amount of benefit and be derived from the Cox proportional hazard model and the decision-making of superiority can be based on the p-value

Table 1 ICH E9(R1) strategies of addressing an intercurrent event

Strategy, as described in the Addendum	Points to consider on selecting the strategy	Considerations on estimation aligned to the strategy
Treatment policy ICE is considered irrelevant in defining the treatment effect of interest; outcome values are used regardless of whether the ICE occurs	This strategy corresponds to a target effect that could be considered most aligned to the effect of treatment assignment (i.e., being prescribed a treatment), but this does not take into account differences between real-world clinical setting and clinical trials If the ICE is related to treatment (such as Tx DC), this strategy definition can be reflected in the Treatment attribute of the estimand (e.g., as in Estimand 1 example) Suggested description of this strategy in the estimand definition: *Strategy targeting the effect of treatment assignment, regardless of the occurrence of this ICE* Becomes meaningless for terminal events, such as death, which might be handled with other strategies such as the Composite Variable strategy	Strategy requires measurements to be collected post-ICE and included in analysis Nonetheless, there will still be missing data in almost any trial and assumptions on the missingness mechanism will need to be made. In general, applying analysis methods for this strategy based on the Missing at Random assumption without accounting for the occurrence of the ICEs expected to change the patient outcome trajectory (such as treatment discontinuation) could lead to bias [10–12]
Hypothetical A scenario is envisaged in which the ICE would not occur; outcome value is that which the variable would have taken in (that) scenario	This strategy could be useful for ICEs that are not considered part of the treatment of interest. An example is when the target effect of assignment to the MDD experimental drug versus placebo does not aim to include the effect of "starting other pharmacological treatments for MDD" in the context of a drug developed as monotherapy treatment The envisaged hypothetical scenario needs to be clearly described in the estimand definition, not left *as if ICE would not occur*. Estimand 1 example description of this strategy is: *A scenario is envisaged in which the event would not have occurred because other pharmacological treatments for MDD are not available*. Estimand 2 example description is: *A scenario is envisaged in which patients would continue treatment as assigned (rather than starting other pharmacological treatments for MDD)*. The Addendum mentions that: "A wide variety of hypothetical scenarios can be envisaged, but some scenarios are likely to be of more clinical or regulatory interest than others." The relevance of each proposed hypothetical scenario needs to be justified If the ICE is related to treatment, this strategy definition can be incorporated in the Treatment attribute of the estimand (e.g., as in Estimand 1 example)	Description of the hypothetical scenario could lead to different assumptions, which could lead to different aligned estimators (see Estimand 1 and 2 examples) When this strategy is used, in most cases data after the ICE is not considered useful in estimation and is therefore not used in analysis. There are also methods that apply adjustments to post-ICE data to model the patient trajectory under the envisaged hypothetical scenario [13, 14] An estimator under the Missing at Random (MAR) such as in example Estimand 2 might be considered to target an effect under optimal conditions (i.e. as if there were no ICEs) and may over-estimate effectiveness in clinical practice
Composite Variable The ICE is considered to be informative and is incorporated into the definition of the variable	This strategy reflects the belief that the ICE itself has a (negative or positive) clinical meaning, such as non-response or relapse (e.g., see Estimands 3 and 5 examples)	Direct incorporation of ICE occurrence in the variable through dichotomization tends to involve loss of information. This may be attenuated by ranking strategies [15], assigning unfavorable values to a continuous outcome [16] or by using post-occurrence multiply-imputed outcome from a suitable alternative distribution [17] Incorporating diverse components may make results difficult to interpret; this might require supplementary analyses exploring effects on composite endpoint components

Table 1 (continued)

Strategy, as described in the Addendum	Points to consider on selecting the strategy	Considerations on estimation aligned to the strategy
While on treatment Outcomes prior to the ICE are of interest	This strategy is applicable to ICEs related to the discontinuation or deviation from the treatment regimen of interest. In the estimand definition, this strategy is reflected in the Variable attribute. Since only pre-ICE measurements are of interest, the Variable cannot relate to a fixed time point, as the duration of study treatment is patient dependent Examples of suitable variables include the area under the curve (see Estimand 7a example), average or a slope derived from pre-ICE measurements or the last measurement prior to or including the ICE occurrence time. This variable must be considered meaningful in the clinical context	This strategy reduces the amount of missing data as the variable is mostly defined based on observed data [18]
Principal stratum The target population is the principal stratum in which an ICE would/would not occur under a certain (set of) treatment assignment(s)	In the estimand definition, this strategy is reflected in the Population attribute. Estimand 4 example has the following definition for Population: *Stratum of patients with a diagnosis of MDD, who would adhere to drug X for 8 weeks (i.e., would comply with, complete the treatment, and would not start other pharmacological treatments for MDD by Week 8, if given drug X)* Could take account of the intercurrent event of death by, e.g., targeting the stratum that would not die while on study, irrespective of randomized treatment group As membership in a stratum cannot be observed for all subjects, supplementary analyses estimating probability of belonging to the stratum should also be provided [19]	Membership in a stratum cannot be directly observed for all study participants (e.g., it cannot be observed in the placebo participants in the Estimand 4 example). Modelling assumptions will be required and sometimes research and experience are insufficient. If predictors of membership of the stratum are omitted, the model for membership of the stratum may be inadequate or biased. Standard model checking measures, such as receiver operating characteristic curves and plots of residuals and predicted values, should be presented to allow the reader of the study report to assess the credibility of the prediction of membership of the stratum

from the log-rank test. Further research [22–24] is currently being done on constructing time to event methods that could be used for both the main estimator and decision-making.

Section "Estimand Examples for Major Depressive Disorder" includes several examples of estimator specifications.

Estimand Examples for Major Depressive Disorder

The ISCTM Estimand WG chose MDD to exemplify the process to select and construct estimand, knowing that:

– It is highly prevalent [25, 26] and extensively studied, with widely accepted endpoints.
– Nevertheless, it is a complex indication to pursue, with many challenges, including high treatment dropout rates.
– Many issues encountered in defining estimands in clinical trials of treatment for MDD can be generalized and applied to clinical trials in many other disease areas. These issues include a relatively high number of discontinuations from treatment, (partial) compliance, and starting other pharmacological treatments for MDD that could influence the trial outcomes.

MDD is defined in the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5-TR) [27], by the occurrence of one or more major depressive episodes. Such episodes must be of at least 2 weeks duration, with at least five of nine specified symptoms co-occurring during that period, not attributable to other causes, and leading to impairment of function compared to a state prior to symptom onset. These episodes comprise a primary symptom of subjective or observed persistence and prevalence of either (1) depressed mood (i.e., sad, empty, or hopeless) or (2) markedly diminished interest or pleasure in almost all activities, and additional potential symptoms of (3) spontaneous loss of appetite or weight, (4) insomnia or hypersomnia, (5) fatigue, (6) observable psychomotor retardation or agitation, (7) impairment in ability to think, concentrate, or make decisions, (8) inappropriate feelings of worthlessness or guilt, and (9) recurrent thoughts of death, particularly suicide.

The symptomatic presentations and durations of episodes, and presence, frequency, and patterns of recurrence, as well as level of subsyndromal inter-episodic symptoms are all highly variable both between and within individuals. Thus, pertinent features of MDD as a clinical entity that may impact the choice of estimand in a clinical trial are:

• No single common pathophysiology—samples may comprise pathophysiologic subpopulations that inform patient strata.

• Episodes may be characterized by multiple symptom dimensions [28]—outcome measures must be appropriately responsive to differential treatment effects on symptom dimensions.
• Typical symptoms may differ depending on patient age (e.g., more negative valence system symptoms in younger adults, more prominent positive valence system deficits in older adults) [28]—such differences may inform selection of outcome measures and characterization of patient strata.
• Episodes can have gradual or abrupt onset and offset and duration ranges widely from a defined minimum of 2 weeks, to over a year [29]—consideration of such features is important for time-based elements of study endpoints.
• Episode duration may also differ depending on patient age [30].
• Episode recurrence rates are variable [29]—consideration of such features is important for time-based elements of study endpoints and relevant ICEs.

For the evaluation of monotherapy treatment, short-term, placebo-controlled trials with or without an active reference arm are the usual standard. The short-term, acute treatment trials are typically followed by long-term, randomized withdrawal trials. Drugs may also be developed to be used as adjunctive treatments to existing antidepressant therapy. The MDD estimand examples in this section are presented in the following type of context:

• Short-term monotherapy MDD treatment
• Maintenance monotherapy MDD treatment
• Short-term adjunctive MDD treatment
• Maintenance adjunctive treatment in patients with treatment resistant MDD (TRD).

The MDD examples included in this section follow the estimand framework steps recommended in Section "Process for Selecting and Constructing Estimands". Some of the examples include specifications for a potential trial design, key trial implementation elements needed to address the estimand, and main and sensitivity estimator specifications that include the elements recommended in Sect. Process for Selecting an Estimator Aligned with an Estimand. It is important to emphasize that the presented estimand and estimator examples are not to be taken as guidance; estimand attributes could be described differently and some of the included elements are subject to further research, especially in the field of aligning estimand and estimators. Each of the five strategies for handling ICEs is addressed in at least one example; all examples are considered to be applicable to MDD, based on the authors' experience.

Estimand 1:

Context	Short-term monotherapy treatment in MDD
Stakeholder	Health Authority Agency
Decision to be made	Determine if the study contributes substantial evidence of short-term efficacy for drug X
Objective	To assess the superiority of drug X versus placebo in short-term symptom reduction when given monotherapy treatment in MDD patients
Intercurrent Events	Tx DC, Starting other pharmacological treatments for MDD
Question of interest	For a patient with MDD for whom acute drug monotherapy would be indicated, what would be the expected effect of prescribing drug X on depression severity at Week 8, were no other antidepressant medications available?
Estimand Definition (The names of attributes in bold are per ICH E9(R1) document and should not be changed.)	**Treatment condition of interest vs Alternative treatment condition:** Assignment to drug X vs placebo, at the selected dose and frequency of administration, regardless of treatment discontinuation and as if other pharmacological treatments for MDD were not available **Population:** Patients with a diagnosis of MDD in a current major depressive episode with at least moderate symptom severity **Variable:** Change from baseline to Week 8 in the total score of the 17-item version of Hamilton Depression Rating Scale (HDRS) [31] **Population-level summary:** Difference in means between treatment conditions **Intercurrent events and Corresponding Strategies:**
Utility of this question of interest and corresponding estimand to Stakeholder	Answering this question requires an estimate of the expected effect of treatment under trial conditions as close as possible to real-world use. The evaluation of the assignment to either drug X vs placebo is of practical importance as treatment discontinuation occurs not only in trials but also in clinical practice and a treatment effect in an ideal condition, where there is perfect compliance, would be unrealistic. On the other hand, the effect of other antidepressants that might be used following assignment to either drug X or placebo is not of interest.
Example: Trial design	Parallel, double-blind (DB), placebo controlled, randomized trial design

Intercurrent events and Corresponding Strategies (within Estimand Definition):

Intercurrent Event	Strategy	Description*
Tx DC	Treatment-policy, as reflected in the Treatment definition	Strategy targeting the effect of treatment assignment, regardless of the occurrence of this ICE
Starting other pharmacological treatments for MDD	Hypothetical, as reflected in the Treatment definition	A scenario is envisaged in which the event would not have occurred because other pharmacological treatments for MDD are not available

* Description of a strategy can be omitted from this table if that strategy is already incorporated into another attribute, as in this example.

Key implementation elements needed to address this estimand	Efficacy data after Tx DC are directly relevant to this estimand and efforts should be made to collect them. Efficacy data after starting other pharmacological treatments for MDD are not relevant to this estimand, but study protocols should plan to collect the information because the data may be relevant to a different or supplementary estimand. The date of occurrence of each ICE should be clearly documented in eCRF.
Estimand and Estimator aligned analysis set	Values of the outcome measure collected from baseline to Week 8 in all randomized participants, **including the values collected after Tx DC**.
Data not used	Values of the outcome measure collected after the ICE of Starting other pharmacological treatments for MDD, addressed by a hypothetical strategy
Missing data	After study withdrawal or missing due to missed visits or missed data collections not related to study withdrawal
Assumptions for data not used and missing	**For Intermittent Missing:** Missing at Random (MAR) assumption, implying that measurements are assumed similar to those from the other participants from same treatment group, who do not have intermittent missing measurements

For Data not Used and Monotone Missing Data: Conditional on observed baseline attributes and on observed post-baseline outcomes, outcomes are assumed similar to those in the placebo group.

Note:

The "similar to placebo" assumption is considered an option under the hypothetical scenario of other pharmacological treatments for MDD not being available and for missing data due to study withdrawal. |
| Main Estimator | • Impute intermittent missing based on Monte Carlo Markov Chains (MCMC), assuming MAR
• Impute data not used and monotone missing data based on the Copy Reference MI method (see Note below)
• Analysis based on Mixed Model for Repeated Measures (MMRM) or Analysis of Covariance (ANCOVA)
• Combine results based on Rubin's rules.

Note:

Commonly used multiple imputation (MI) estimators [17, 32] using a "similar to placebo" assumption are 1) "Copy Reference", where the model-expected mean for imputations is derived as if the participants had always been a member of the placebo group; and 2) "Jump to Reference", where the model-expected mean for imputations is derived relative to the subject's own treatment group, but then shifted so as to be relative to the control group for the imputed visit. Under this approach, participants from the drug X group are considered as not being treated any longer (and therefore similar to placebo participants) and participants from the placebo group are considered to have trajectories similar to those from their own treatment arm who remained on treatment.

Conditioning on baseline and post-baseline data:

Under variant 1), insofar as a subject's observed outcomes are better/worse *than those of subjects in the control group*, the mean for imputations for that subject at the imputed visit will tend to be better/worse than those observed for the control group at the imputed visit.

Under variant 2), insofar as a subject's observed outcomes are better/worse *than those of subjects in the subject's own treatment group*, the mean for imputations for that subject at the imputed visit will tend to be better/worse than those observed for the control group at that visit. |

Sensitivity Estimator(s), including what assumptions change from the main estimator	**Sensitivity Estimator:** Same steps as in the Main Estimator, except that in the second step a Jump to Reference MI method is used instead. **Change in the assumptions for data not used or missing**: See Note from Main Estimator.
Analysis used for decision making, mentioning if different or same as the Main estimator	Same as Main estimator
Other comments	None

Note: A different main estimator for same estimand can be constructed using different

assumptions (see underlined text), as exemplified below:

Assumptions for data not used and missing	**For Intermittent Missing:** MAR assumption, implying that measurements are assumed similar to those from the other participants from same treatment group, who do not have intermittent missing measurements. **For Data not Used and Monotone Missing Data:** <u>MAR assumption given baseline value, post-baseline observed measurements, treatment group and treatment discontinuation.</u>
Main Estimator	• Impute intermittent missing based on MCMC, assuming MAR • Impute monotone missing and data not used based on a <u>MAR MI method [11] with an imputation model consistent with the assumption described above</u>. • Analysis based on MMRM or ANCOVA • Combine results based on Rubin's rules.
Sensitivity Estimator(s), including what assumptions change from the main estimator	**Sensitivity Estimator: Delta adjustment sensitivity analysis** Same steps as in the Main Estimator, except that in the second step a sequence of negative and positive adjustments (called delta adjustments [17]) is applied to the imputed values after study withdrawal and after the second ICE of starting other pharmacological treatments for MDD. A two-dimensional tipping point sensitivity map is then created, showing the analysis results for all delta adjustments for drug X group vs placebo, highlighting the cases with negative results (e.g. statistically non-significant). Note: A robust sensitivity analysis would show negative results only for the application of severe, clinically not plausible delta adjustments. **Change in the assumptions for data not used or missing:** A Missing Not at Random (MNAR) set-up is applied by assuming that patients are better or worse after study withdrawal and the second ICE than similar patients who don't experience these events.

Estimand 2:

Context	Short-term monotherapy treatment in MDD
Stakeholder	Pharmaceutical company for internal decision-making, such as for a Phase 2 trial. This estimand may also be of interest to prescribers, patients and, in some cases, to a Health Authority Agency (see comments at the end of this estimand section).
Decision to be made	Determine if the study provides enough evidence of efficacy for drug X to decide on continuing its development
Objective	To assess the superiority of drug X versus placebo on symptom reduction when given as short-term monotherapy treatment in MDD patients
Intercurrent Events	Tx DC, Starting other pharmacological treatments for MDD
Question of interest	For a patient with MDD for whom acute drug monotherapy would be indicated, what would be the expected effect of drug X on depression severity at Week 8, if taken as directed for the entire 8 weeks without initiating other MDD treatments?
Estimand Definition (The names of attributes in bold are per ICH E9(R1) document and should not be changed.)	**Treatment condition of interest vs Alternative treatment condition:** Drug X vs Placebo, at the selected dose and frequency of administration, as if patients would continue treatment as assigned (rather than discontinuing investigational treatment or starting other pharmacological treatments for MDD) **Population:** Patients with a diagnosis of MDD in a current major depressive episode with at least moderate symptom severity **Variable:** Change from baseline to Week 8 in the total score of the 17-item version of HDRS **Population-level summary:** Difference in means between treatment conditions **Intercurrent events and Corresponding Strategies:** _(see table below)_ * Description of a strategy can be omitted from this table if that strategy is already incorporated into another attribute, as in this example.
Utility of this question of interest and corresponding estimand to Stakeholder	Answering this question requires an estimate of the expected effect of treatment under optimal trial conditions (i.e. as if there were no ICEs). While this will provide evidence for treatment benefit, additional questions would also need to be addressed, such as on estimating the rate of ICEs by treatment group.
Example: Trial design	Parallel, DB, placebo controlled, randomized trial design

Intercurrent Event	Strategy	Description*
Tx DC	Hypothetical, as reflected in the Treatment definition	A scenario is envisaged in which patients would continue treatment as assigned (rather than discontinuing treatment)
Starting other pharmacological treatments for MDD	Hypothetical, as reflected in the Treatment definition	A scenario is envisaged in which patients would continue treatment as assigned (rather than starting other pharmacological treatments for MDD)

Key implementation elements needed to address this estimand	The date of occurrence of each ICE should be clearly documented in eCRF.
Estimand and Estimator aligned analysis set	Values of the outcome measure collected from baseline to Week 8 in all randomized participants.
Data not used	Values of the outcome measure collected after the ICEs of Tx DC and Starting other pharmacological treatments for MDD, addressed by a hypothetical strategy
Missing data	Missing due to missed visits or missed data collections not related to study withdrawal, such as intermittent missing.
Assumptions for data not used and missing	MAR assumption given treatment assignment, baseline variables and post-baseline variable values (i.e., changes in HDRS total scores), implying that measurements are assumed similar to those from the other participants from same treatment group, who do not experience the ICEs.
Main Estimator	MMRM
Sensitivity Estimator(s), including what assumptions change from the main estimator	**Sensitivity Estimator: Delta adjustment sensitivity analysis** Same as the sensitivity analysis described under Estimand 1, alternative estimator example.
Analysis used for decision making, mentioning if different or same as the Main estimator	Same as Main estimator
Other comments	This estimand is sensitive to the **existence** of a drug effect because it provides an upper bound of the treatment benefit. However, it may not allow for balanced comparison of risks and benefits.

Estimand 3:

Context	Short-term monotherapy treatment in MDD
Stakeholder	Payers This estimand may also be of interest to a Health Authority Agency, prescribers and patients
Decision to be made	Determine if the study contributes substantial/strong evidence of clinically meaningful patient level short-term benefit for drug X
Objective	To assess the extent of clinically meaningful benefit of drug X versus placebo when given as short-term monotherapy treatment in MDD patients
Intercurrent Events	Tx DC, Starting other pharmacological treatments for MDD
Question of interest	For a patient with MDD for whom acute drug monotherapy would be indicated, what would be the expected effect of prescribing drug X on the likelihood of experiencing a treatment response at Week 8? Treatment response requires a substantial improvement of symptoms without starting a different MDD treatment, with no premature treatment discontinuation due to side effects or lack of efficacy.
Estimand Definition (The names of attributes in bold are per ICH E9(R1) document and should not be changed.)	**Treatment condition of interest vs Alternative treatment condition:** Assignment to drug X vs placebo, at the selected dose and frequency of administration, as if patients would not discontinue investigational treatment due to other reasons than AEs and LOE **Population:** Patients with a diagnosis of MDD in a current major depressive episode with at least moderate symptom severity **Variable:** Binary responder variable, where a responder is defined as a participant who has at least 50% reduction from baseline to Week 8 in the HDRS total score and does not discontinue treatment due to AEs and LOE or start other pharmacological treatment for MDD by Week 8 **Population-level summary:** Difference in responder proportions between treatment conditions **Intercurrent events and Corresponding Strategies:** (see table below)
Utility of this question of interest and corresponding estimand to Stakeholder	Answering this question requires an estimate of the expected effect of treatment that can be translated into benefit at the patient level. The definition of **benefit** needs to be relevant and acceptable to the Stakeholder.

Intercurrent Event	Strategy	Description*
Tx DC due to AEs and LOE	Composite Variable, as reflected in the Variable definition	A participant experiencing this ICE is considered a non-responder
Tx DC due to other reasons than AEs and LOE	Hypothetical, as reflected in the Treatment definition	A scenario is envisaged in which patients would not discontinue investigational treatment due to other reasons than AEs and LOE
Starting other pharmacological treatments for MDD	Composite Variable, as reflected in the Variable definition	A participant experiencing this ICE is considered a non-responder

* Description of a strategy can be omitted from this table if that strategy is already incorporated into another attribute, as in this example.

Example: Trial design	Parallel, DB, placebo controlled, randomized trial design
Key implementation elements needed to address this estimand	The date of occurrence of each ICE should be clearly documented in eCRF.
Estimand and Estimator aligned analysis set	Values of the outcome measure collected from baseline to Week 8 in all randomized participants.
Data not used	Values of the outcome measure collected after the ICE of Tx DC due to other reasons than AEs and LOE
Missing data	Values for participants with no ICEs but with a missing Week 8 HDRS total score value, if applicable.
Assumptions for data not used and missing	**For Intermittent Missing:** MAR assumption **For Data not Used and Monotone Missing Data:** MAR assumption given baseline HDRS total score value, post-baseline observed values (i.e., changes in HDRS total scores), treatment group and indicators for Tx DC due to AE and LOE, and Starting other pharmacological treatments for MDD.

Main Estimator	• Impute change in HDRS total score as following:
	    ○ Impute intermittent missing based on MCMC, assuming MAR
	    ○ Impute monotone missing and data not used based on a MAR MI method with an imputation model consistent with the assumption described above, using MI regression.
	• For each of the multiply imputed datasets, compute the responder status for all participants included in the analysis set (apply the Variable definition from the estimand, taking into account that participants with ICEs addressed with the composite strategy are considered non-responders)
	• Compute the mean of the responder proportions across the multiply imputed datasets for each treatment group
	• Compute the difference vs placebo for the means of the responder proportions
	Note: If the binary variable is derived from a continuous one, it is recommended [33] to impute first the continuous variable and then derive the binary value from the imputed values.

Sensitivity Estimator(s), including what assumptions change from the main estimator	**Sensitivity Estimator: Delta adjustment sensitivity analysis** Same steps as in the Main Estimator, except adding an additional step 1c: a sequence of worsening adjustments (called delta adjustments) is applied to the imputed values. A two-dimensional tipping point sensitivity map is created, showing the analysis results for all delta adjustments for drug X group vs placebo, highlighting the cases with negative results (e.g. statistically non-significant, see section on Analysis used for Decision Making). Note: A robust sensitivity analysis would show negative results only for the application of severe, clinically not plausible delta adjustments. **Change in the assumptions for data not used or missing:** A MNAR set-up is applied by assuming a worsening depression severity (so an increasing non-responder status) for the imputed values relative to the MAR estimates.
Analysis used for decision making, mentioning if different than the Main estimator	• Apply first 2 steps from main estimator • Apply to each multiply imputed dataset a Cochran-Mantel-Haenszel (CMH) test, controlling for any stratification factors • Combine results based on Rubin's rules, applying first the Wilson-Hilferty transformation for asymptotic normality • Compute the p-value based on the combined test statistic

| | A two-dimensional tipping point sensitivity map is also created, highlighting the cases with negative results (e.g. see section on Sensitivity Estimators). |
| Other comments | In this example a binary responder variable is defined based on a continuous variable and the occurrence of certain ICEs. Dichotomizing a continuous variable usually leads to loss of used information and hence potential decrease in power. Because of this, estimands that employ a continuous variable instead of a binary version are often preferred for a primary estimand. However, this type of estimand is important in assessing benefit at the patient level and is often used as supplementary, to provide a different way of defining the treatment effect.

Of note, multiple responder criteria could be of interest. Plots of cumulative responder curves can also be created, with the percent reduction from baseline to Week 8 in the HDRS total score used in the responder criteria varying from 0 to 100%. |

Estimand 4:

Context	Short-term monotherapy treatment in MDD
Stakeholder	Prescriber or patient May also be of interest to a Health Authority Agency, payers
Decision to be made	Determine if the study provides sufficient evidence of efficacy to prescribe drug X, by understanding the treatment effect in patients who would adhere to drug X for a certain duration.
Objective	To assess the superiority of drug X vs placebo on symptom reduction in the patients who would adhere to drug X for a certain duration.
Intercurrent Events	Tx DC if assigned to drug X Starting other pharmacological treatments for MDD if assigned to drug X Severe treatment non-compliance (i.e., severe intermittent or partial treatment adherence) if assigned to drug X
Question of interest	For the type of patient with MDD for whom acute drug monotherapy would be indicated, who would take drug X as prescribed and would not initiate other MDD treatments for 8 weeks, what would be the expected effect of drug X on depression severity at Week 8?
Estimand Definition (The names of attributes in bold are per ICH E9(R1) document and should not be changed.)	**Treatment condition of interest vs Alternative treatment condition:** Assignment to drug X vs placebo, at the selected dose and frequency of administration **Population:** Stratum of patients with a diagnosis of MDD in a current major depressive episode with at least moderate symptom severity, who would adhere to drug X for 8 weeks (i.e., would comply with, complete the treatment and would not start other pharmacological treatments for MDD by Week 8, if given drug X) **Variable:** Change from baseline to Week 8 in the total score of the 17-item version of the HDRS **Population-level summary:** Difference in means between treatment conditions **Intercurrent events and Corresponding Strategies:** <table><tr><th>Intercurrent Event</th><th>Strategy</th></tr><tr><td>**Tx DC if assigned to drug X**</td><td>Principal Stratum, as reflected in the Population definition</td></tr><tr><td>**Starting other pharmacological treatments for MDD if assigned to drug X**</td><td>Principal Stratum, as reflected in the Population definition</td></tr><tr><td>**Severe treatment non-compliance (i.e., severe intermittent or partial treatment adherence) if assigned to drug X**</td><td>Principal Stratum, as reflected in the Population definition</td></tr></table> * Description of a strategy can be omitted from this table if that strategy is already incorporated into another attribute, as in this example.
Utility of this question of interest and corresponding	Estimating treatment effect in the stratum of patients who would adhere to drug X for its intended duration, complemented by estimands/questions related to non-adherers due to different reasons

estimand to Stakeholder	(e.g., due to AE or lack of efficacy), could be of interest to a variety of stakeholders [19].
Example: Trial design	Parallel, DB, placebo controlled, randomized trial design
Key implementation elements needed to address this estimand	Comprehensive collection of baseline variables potentially predicting ICEs. The date of occurrence of each ICE should be clearly documented in eCRF.
Estimand and Estimator aligned analysis set	For all randomized participants, values of the outcome measure collected from baseline to Week 8 for participants with no ICEs and from baseline to ICE for participants with ICEs
Data not used	Values of the outcome measure after any of the ICEs in either treatment group
Missing data	Missing due to missed visits or missed data collections not related to study withdrawal, such as intermittent missing.
Data not existing	For participants not assigned to drug X, data on the occurrence of ICEs had they been assigned to drug X instead
Assumptions for data not used, not existing and missing	Probability of being an adherer under drug X can be estimated from the pre-determined baseline variables. Outcomes in the placebo group are clinically informative and provide a meaningful comparison vs. drug X experimental group, including when participants assigned to placebo experience ICEs.
Main Estimator	Estimator defined by the following steps (for implementation, see Estimating Principal Strata, site of the Drug Information Association Scientific Working Group on Estimands and Missing Data [34]): - Define a MI model for presence/absence of ICEs based on the observed ICEs (response) and the pre-determined baseline characteristics from drug X group (explanatory variables). - For each participant in the placebo group, multiply impute M times the presence/absence of ICEs had the participant been assigned to drug X instead, using the participant's observed baseline characteristics. - In each of the M imputed datasets, select the participants in the principal stratum of patients not experiencing ICEs under drug X. They include the participants from the drug X group who did not experience ICEs and the participants from the placebo group who were imputed not having any ICEs had they been assigned to drug X instead. - Estimate treatment difference in each of the M principal strata (apply MMRM) and combine results using Rubin's rules. Notes: Standard model testing measures will be used to assess the bias and variance of the MI model for ICE presence/absence based on the experimental group. Average number of subjects modelled as present in the principal stratum will be presented by treatment group. Summary statistics for the model explanatory variables of subjects from the principal stratum will also be presented by treatment group.
Sensitivity Estimator(s), including what assumptions change from the main estimator	Options include defining MI models using a different set of baseline characteristics (e.g. a subset of the initial set) or using alternatives analysis methods than MMRM. Alternative estimators (based on different assumptions) can also be used [35,36].

Analysis used for decision making, mentioning if different than the Main estimator	Same as main estimator
Other comments	A randomized withdrawal trial design with a drug X run-in period, after which all adherers/responders are randomized to either continue drug X or switch to placebo, could also be considered to evaluate treatment effect in drug X adherers/responders. However, this design addresses a different question of interest (see Estimand 5) and also comes with other potential concerns such as withdrawal effect when drug X is switched to placebo.

Estimand 5:

Context	Maintenance monotherapy treatment for MDD
Stakeholder	Health Authority Agency
Decision to be made	Determine if the study provides substantial evidence that drug X is effective as a maintenance treatment for MDD after an initial short-term response
Objective	To assess the superiority of drug X versus placebo in preventing relapse in patients with MDD who have shown a stable response to initial treatment with drug X and for whom continuing monotherapy would be clinically acceptable.
Intercurrent Events	Tx DC due to reasons other than those included in the definition of relapse
Question of interest	For a patient with MDD who experienced a stable response to initial treatment with drug X and for whom continuing monotherapy would be clinically acceptable, what is the effect of continuing versus discontinuing drug X on the occurrence of relapse up to 1 year?
Estimand Definition (The names of attributes in bold are per ICH E9(R1) document and should not be changed.)	**Treatment condition of interest vs Alternative treatment condition:** Assignment to continuation of drug X vs switching to placebo, at the selected dose and frequency of administration, regardless of treatment discontinuation due to other reasons than included in the definition of relapse **Population:** Patients with a diagnosis of MDD, who have shown a stable response to initial treatment with drug X and for whom continuing monotherapy would be clinically acceptable **Variable:** Time to relapse up to Year 1, where relapse is defined as the first occurrence of 1) a total score >X on the 17-item version of the HDRS, 2) hospitalization due to depressive or MDD-associated symptoms (including suicidal ideation or behavior), 3) treatment discontinuation due to lack of efficacy and/or suicidal ideation or behavior, 4) switching to or adding other pharmacological treatment for MDD **Population-level summary:** Hazard ratio of drug X versus placebo **Intercurrent events and Corresponding Strategies:** table below
	Intercurrent Event | **Strategy** | **Description***

Intercurrent Event	Strategy	Description*
Tx DC due to reasons other than those included in the definition of relapse	Treatment-policy, as reflected in the Treatment definition	Strategy targeting the effect of treatment assignment, regardless of the occurrence of this ICE

	* Description of a strategy can be omitted from this table if that strategy is already incorporated into another attribute, as in this example.
Utility of this question of interest and corresponding estimand to Stakeholder	The evaluation of the assignment to continuation of drug X versus switching to placebo is of practical importance as decisions on treatment continuation in remitted patients need to be informed.
Example: Trial design	Randomized withdrawal trial design, where patients who can be stabilized on drug X in a run-in open-label phase are then randomized into a parallel, DB, placebo-controlled phase, comparing stabilized dose of drug X to placebo
Key implementation elements needed to address this estimand	Efficacy and hospitalization data after Tx DC due to other reasons than included in the definition of relapse are directly relevant to this estimand as occurrence of relapse can still be assessed, and efforts should be made to collect them. Other ICEs are included in the definition of the primary outcome variable. The date of occurrence of each ICE should be clearly documented in eCRF.

Estimand and Estimator aligned analysis set	Relapse based on outcome values collected from randomization to Year 1 in all randomized participants who receive at least one dose of study drug during the DB phase, including the relapses collected after Tx DC due to other reasons than included in the definition of relapse.
Data not used	Outcome values collected after occurrence of relapse are not used in the analysis for this estimand as the endpoint has been reached and further data are not relevant.
Missing data	After study withdrawal
Assumptions for missing data	Censoring applied at study withdrawal is assumed ignorable. Censoring at termination of the trial is also assumed ignorable (administrative censoring).
Main Estimator	Cox-regression (using proportional hazard assumption)
Sensitivity Estimator(s), including what assumptions change from the main estimator	**Sensitivity Estimator: Delta adjustment sensitivity analysis** The main analysis relies on the assumption of ignorable censoring. Therefore, sensitivity analyses will be performed to stress-test the robustness of results to deviations from ignorable censoring. Specifically, it is assumed in this sensitivity analysis that subjects who withdraw from the study with no recorded relapse have a higher relapse hazard starting from the study withdrawal (SW) time, compared with similar subjects who remain in the study. The higher relapse hazard is determined by the sensitivity parameter Delta, representing the ratio of subject-specific hazard at any given time point t following SW compared to that same subject's hazard at the same time t if he or she had remained in the study. A semi-parametric multiple imputation approach will be used for the imputation of relapse events, as described in Lipkovich et al [37]. A two-dimensional tipping point sensitivity map will be created, showing the analysis results for all Delta adjustments for drug X group vs placebo, highlighting the cases with negative results (e.g., statistically non-significant). Note: A robust sensitivity analysis would show negative results only for the application of severe, clinically not plausible delta adjustments. **Change in assumptions:** Censoring applied at study withdrawal is assumed *non-ignorable*. Note: Another sensitivity estimator can also be proposed that does not rely on the proportional hazard assumption [38].
Analysis used for decision making, mentioning if different or same as the Main estimator	Same as Main estimator
Other comments	Potential occurrence of withdrawal-syndrome-induced relapse after discontinuing drug X is one of the concerns of this estimand and corresponding design. Estimand 5a, which could be used as supplementary, attempts to mitigate this concern. Under Estimand 5, the trial should be designed in such a way as to minimize concern of acute and subacute withdrawal effects – e.g., minimum duration of treatment in the run-in phase to establish responder status and pharmacologically appropriate tapering regimen (to minimize possibility of acute withdrawal effects). Supplementary analyses could be considered including all-cause Tx DC in the relapse definition (so applying a composite strategy to all considered ICEs) or applying a treatment policy strategy to all considered ICEs.

Estimand 5-alt:

The alternative estimand 5-alt attempts to mitigate the concern of potential occurrence of "withdrawal-syndrome-induced relapse" after discontinuing drug X. The fields in this example are same as in Estimand 5, unless presented in the table below.

Context	Maintenance monotherapy treatment for MDD
Stakeholder	Health Authority Agency
Decision to be made	Determine if the study provides substantial evidence that drug X is effective as a maintenance treatment for MDD after an initial short-term response
Objective	To assess the superiority of drug X versus placebo in preventing relapse in patients with MDD who have shown a stable response to initial treatment with drug X, for whom continuing monotherapy would be clinically acceptable *and who, if discontinued from drug X, would not have a withdrawal-syndrome-induced-relapse*.
Question of interest	For a patient with MDD who experienced a stable response to initial treatment with drug X, for whom continuing monotherapy would be clinically acceptable *and who, if discontinued from drug X, would not have a withdrawal-syndrome-induced-relapse*, what is the expected effect be of continuing versus discontinuing drug X on the occurrence of relapse up to 1 year?
Estimand Definition (The names of attributes in bold are per ICH E9(R1) document and should not be changed.)	**Note: Estimand attributes are same as in Estimand 5 unless mentioned otherwise.** **Population:** Patients with a diagnosis of MDD, who have shown a stable response to initial treatment with drug X *and who, if discontinued from Drug X, would not have a withdrawal-syndrome-induced-relapse* **Variable:** *Same as for Estimand 5, except that the events considered "withdrawal-syndrome-induced-relapse" are not applied towards the relapse definition* **Intercurrent events and Corresponding Strategies:**
Utility of this question of interest and corresponding estimand to Stakeholder	The evaluation of the assignment to continuation of drug X versus switching to placebo in patients *who, if switched to placebo, would not have a withdrawal-syndrome-induced relapse* is of practical importance as decisions on treatment continuation in remitted patients need to be informed and the risk of relapse due to discontinuation of treatment quantified separately from the risk of relapse due to the withdrawal effect, and also because the initial decision to treat should be informed by all the long-term consequences of such decision.

Intercurrent Event	Strategy	Description*
Tx DC due to other reasons than included in the definition of relapse	Treatment-policy, as reflected in the Treatment definition	Strategy targeting the effect of treatment assignment, regardless of the occurrence of this ICE
Withdrawal-induced-relapses	Principal stratum, as reflected in the Population definition	

* Description of a strategy can be omitted from this table if that strategy is already incorporated into another attribute, as in this example.

Comments	The population is defined under the principal stratum strategy as patients who, had they have been assigned to discontinuing the investigational drug, would not experience withdrawal-syndrome-induced relapses. Assuming that criteria to classify events as "withdrawal-syndrome-induced relapses" rather than as outcome-events can be defined a priori (for example, based on timing plus other phenotypical features), those criteria can be applied to the events that are observed in the patients randomized to discontinue the investigational drug. Hence, for patients who are assigned to discontinue the investigational drug, the pertinence to the stratum of interest can be directly observed (i.e., it can be observed whether they did not encounter withdrawal-syndrome-induced-relapses). Comparing this subset of patients assigned to discontinue the investigational drug with all patients not assigned to discontinue the investigational drug (for which, that is, the pertinence to the stratum of interest cannot be observed) would possibly result in a biased comparison. One way to correct for this would be to build on the observed withdrawal-syndrome-induced relapses in the patients assigned to discontinuing the investigational drug a propensity score (based on observed baseline and post-baseline information) that can be applied to correct the analysis. As in Estimand 4, this approach assumes that the probability of the event (i.e., the probability of having a withdrawal-syndrome-induced relapse) can be estimated from pre-determined baseline variables).

Estimand 6:

Context	Short-term add-on/adjunctive treatment in MDD
Primary Stakeholder	Health Authority Agency
Decision to be made	Determine if the study contributes substantial evidence of efficacy for drug X as short-term add-on treatment to the underlying class Y of antidepressant treatments (ADT) in MDD patients who have had an inadequate response to current ADT
Objective	To assess the superiority of drug X versus placebo on symptom reduction, when given as short-term add-on treatment to the underlying class Y of ADT in MDD patients who have had an inadequate response to current ADT
Intercurrent Events	Tx DC of add-on investigational drug X, Tx DC of the underlying ADT, starting other add-on and/or underlying pharmacological treatment for MDD
Question of interest	For a patient with MDD who has had an inadequate response to their current ADT, what would be the expected effect on depression severity at Week 8 of prescribing the addition of drug X, were no other ADTs available?
Estimand Definition (The names of attributes in bold are per ICH E9(R1) document and should not be changed.)	**Treatment condition of interest vs Alternative treatment condition:** Assignment to the add-on drug X vs placebo, taken together with an underlying ADT from class Y, at the selected dose and frequency of administration, regardless of add-on treatment or underlying ADT discontinuation and as if other pharmacological treatments for MDD were not available **Population:** Patients with a diagnosis of MDD in a current major depressive episode with at least moderate symptom severity, who have had an inadequate response to current ADT **Variable:** Change from baseline to Week 8 in the total score of the 17-item version of the HDRS **Population-level summary:** Difference in means between treatment conditions **Intercurrent events and Corresponding Strategies:**

Intercurrent Event	Strategy	Description*
Tx DC of add-on investigational drug	Treatment-policy, as reflected in the Treatment definition	Strategy targeting the effect of treatment assignment, regardless of the occurrence of this ICE
Tx DC of the underlying ADT	Treatment-policy, as reflected in the Treatment definition	Strategy targeting the effect of treatment assignment, regardless of the occurrence of this ICE
Starting other add-on and/or underlying pharmacological treatment for MDD	Hypothetical, as reflected in the Treatment definition	A scenario is envisaged in which the event would not have occurred because other pharmacological

		treatments for MDD are not available
	* Description of a strategy can be omitted from this table if that strategy is already incorporated into another attribute, as in this example. Note: As drug X is being evaluated **only** as add-on treatment to the underlying class Y of ADT and it is not intended as monotherapy, there could be situations in which if patients discontinue the initial underlying ADT then they are advised to immediately switch to a different one. In these situations, Tx DC of the underlying ADT will not be considered as a stand-alone ICE as it will be covered by the last ICE.	
Utility of this question of interest and corresponding estimand to Stakeholder	The evaluation of the assignment to the add-on drug X vs placebo, regardless of add-on treatment or underlying ADT discontinuation, is of practical importance as treatment discontinuation occurs not only in trials but also in clinical practice and a treatment effect in an ideal condition where there is perfect compliance would be unrealistic. In addition, the effect of other antidepressants that patients might switch to is not of interest.	
Example: Trial design	Parallel, DB, placebo controlled, randomized trial design	
Key implementation elements needed to address this estimand	Efficacy data after Tx DC of add-on treatment or underlying ADT are directly relevant to this estimand and efforts should be made to collect them. The date of occurrence of each ICE should be clearly documented in eCRF.	
Estimand and Estimator aligned analysis set	Values of the outcome measure collected from baseline to Week 8 in all randomized participants, **including the values collected after Tx DC of add-on treatment or underlying ADT**	
Data not used	Values of the outcome measure collected after the ICE of Switch of add-on treatment and/or underlying ADT to other pharmacological treatments, addressed by a hypothetical strategy	
Missing data	After study withdrawal or missing due to missed visits or missed data collections not related to study withdrawal	
Main and Sensitivity Estimators	See Estimand 1	

Estimands 7a and 7b:

The following estimand examples from the context of maintenance add-on/adjunctive treatment in MDD were inspired by the LQD study description from Marwood et al. [39]. They do not reflect exactly this trial original objectives and are provided as an example of estimands that complement each other. As a different example from same context, an estimand that could be aligned with the randomized withdrawal trial presented in Brunner et al. [40] could have common elements with Estimand 5 so it has not been used as an additional example for this manuscript.

Estimands 7a and 7b, defined in the following, could either be considered co-primary estimands (if the objective is to show superiority on both) or one could be considered primary and the other supplementary.

	Estimand 7a	Estimand 7b
Context	Maintenance add-on/adjunctive treatment in patients with treatment resistant MDD (TRD)	
Primary Stakeholder	Payers May also be of interest to prescribers and patients.	
Decision to be made	Determine whether the decision to prescribe drug X is more clinically effective over a long-term period than the decision to prescribe drug Y.	
Objective	To assess the superiority of add-on drug X versus add-on drug Y to underlying ADT, when given as maintenance add-on treatment in patients with TRD.	
Intercurrent Events	Tx DC of the add-on drug; Other treatment regimen modifications (Tx DC of the underlying ADT, switch of the underlying ADT or addition of any other concomitant pharmacological and non-pharmacological Interventions)	
Question of interest	For a patient with TRD, what symptom burden can be expected after assignment to Drug X versus Drug Y as add-on treatment to underlying ADT, while the add-on treatment is being taken up to 1 year and regardless of other treatment regimen modifications?	How long can patients with TRD be expected to comply with an add-on treatment with Drug X versus Drug Y, regardless of other treatment regimen modifications?
Estimand Definition (The names of attributes in bold are per ICH E9(R1) document and should not be changed.)	**Treatment condition of interest vs Alternative treatment condition:** Assignment to drug X vs drug Y as add-on treatment to underlying ADT, regardless of other treatment regimen modifications (see definition of this ICE); **Population:** Patients with a diagnosis of TRD and need for additional treatment (irrespective of symptom severity); **Variable:** Standardized area under the curve (AUC) based on the self-rated QIDS-SR values up to Year 1 or Tx DC of the add-on drug, whichever occurs first, defined as the AUC divided by the duration on treatment **Population-level summary:** Difference in means between drug X and drug Y	**Treatment conditions** and **population** attributes as in Estimand 7a **Variable:** Time to Tx DC of the add-on drug from the time of first prescription **Population-level summary:** Difference at Year 1 in the Restricted Mean Survival Time (RMST) values between drug X and drug Y

Estimand 7a — Intercurrent events and Corresponding Strategies:

Intercurrent Event	Strategy	Description*
Tx DC of the add-on drug	While-on-treatment, as reflected in the Variable definition	Strategy targeting a treatment effect captured while the add-on treatment is being taken
Other treatment regimen modifications	Treatment-policy, as reflected in the Treatment definition	Strategy targeting the effect of decision to prescribe drug X vs drug Y as add-on treatment, regardless of the occurrence of this ICE

* Description of a strategy can be omitted from this table if that strategy is already incorporated into another attribute, as in this example.

Estimand 7b — Intercurrent events and Corresponding Strategies:

Intercurrent Event	Strategy	Description*
Other treatment regimen modifications	Treatment-policy, as reflected in the Treatment definition	Strategy targeting the effect of decision to prescribe drug X vs drug Y as add-on treatment, regardless of the occurrence of this ICE

* Description of a strategy can be omitted from this table if that strategy is already incorporated into another attribute, as in this example.

Utility of this question of interest and corresponding estimand to Stakeholder	The evaluation of the decision to prescribe add-on drug X vs drug Y on long-term outcomes, regardless of any changes in the underlying ADT treatment, is of practical importance as it can inform the use of effective options in clinical practice and treatment reimbursement.
Comments	The standardized AUC used as a variable in Estimand 7a provides a "while-on-treatment" measure of the treatment effect that takes into account different durations on treatment, giving an average score while on treatment. Estimand 7b, which uses the duration on treatment as the variable, complements Estimand 7a by providing a different facet of the treatment effect. The Addendum Section A.3.4. mentions that ". . . an estimand using a while on treatment strategy should usually be accompanied by the additional information on the time to intercurrent event distributions". Use of the While-on-treatment strategy for the ICE of Tx DC of the add-on drug is not implied in the Marwood et al but could be considered a potential option for Estimand 7a. While it is difficult to reconstruct an estimand from a paper that did not use the estimand framework, the Statistical Methods section suggests that a treatment policy strategy would have been used for all ICEs: "The main analysis will follow an intention to treat (ITT) principle, whereby patient data are analyzed by treatment group, regardless of the medication status of patients throughout the follow-up period."
Example: Trial design	Parallel group, multi-center, pragmatic, open label, patient randomized clinical trial

Discussion

This paper describes an interdisciplinary process for implementing the estimand framework proposed by the ISCTM Estimand WG, a group that represents both clinical and statistical functions. Building on Bell et al. [41] and Ratitch et al. [42, 43], we expand the "thinking process" outlined in the ICH E9(R1) official training material [44] by considering the trial stakeholder(s), the decisions they need to make and the questions that would support their decision making. Study teams are encouraged to justify how answering the proposed questions of interest would support stakeholder decision-making.

The thinking process proposed is reflected in multiple examples using hypothetical trials evaluating a treatment for MDD. While this process is relevant to any therapeutic setting, all examples have been chosen to be applicable to this disease state, based on the authors' experience.

While multiple estimand examples have been included for a given context, such as short-term monotherapy treatment in MDD, each example followed the recommended process, with clarity on the stakeholder, the decision to be made and the corresponding objective and question of interest. This is different from the previous practice (that the Addendum aims to curtail) of running multiple "sensitivity analyses", without thought to what they estimate and their usefulness and purpose. With regard to sensitivity analyses, the Addendum recommends instead a structured approach to stress-test the assumption of the main estimator. This has been reflected in the sensitivity analyses exemplified in this paper.

In this paper we focus on the process of defining the estimand itself and do not directly address in detail the implications for the study procedures. However, the defined estimands will be reflected in the design of a study, from consent form through duration and level of follow-up to final analysis. For example, we note that selecting the estimand will lead the study team to consider logistical elements of study including.

- the burden of the study for participants (the duration of follow-up, the number of visits, complexity of data collection)
- whether to continue follow-up after an ICE (e.g., possibility of subjects remaining in the study after ICEs such as discontinuation of study treatment)
- flexibility to collect some but not all protocol assessments after treatment discontinuation or other ICE

Ultimately this paper highlights the need to incorporate multi-disciplinary collaborations into implementing the ICH E9(R1) framework and provides extensive examples on how this can be accomplished. The process described includes the element of estimand justification to foster alignment within study teams, to ensure that trials will provide answers to the most relevant clinical questions for key trial stakeholders.

References

1. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). Updated Nov 20 2019. https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf. Accessed Sept 7 2022

2. Protocol template for phase 2 and 3 clinical trials that require FDA-IND or IDE application. National Institutes of Health (NIH). Updated Apr 7 2017. https://grants.nih.gov/policy/clinical-trials/protocol-template.htm. Accessed Sept 7 2022

3. Common Protocol Template (CPT). TransCelerate BioPharma INC, Clinical Content & Reuse Solutions. Updated 2021. https://www.transceleratebiopharmainc.com/assets/clinical-content-reuse-solutions/. Accessed Sept 7 2022

4. Chronic Rhinosinusitis with Nasal Polyps: Developing Drugs for Treatment, Guidance for Industry. U.S. Food and Drug Administration (FDA). Updated Dec 16 2021. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/chronic-rhinosinusitis-nasal-polyps-developing-drugs-treatment. Accessed Feb 27 2023

5. Guideline on the clinical investigation of medicines for the treatment of Alzheimer's disease. European Medicines Agency (EMA). Updated Sept 1 2018. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicines-treatment-alzheimers-disease-revision-2_en.pdf. Accessed Mar 7 2023

6. Guidance to Sponsors on How to Manage Clinical Trials During the COVID-19 Pandemic. European Medicines Agency Committee for Medicinal Products for Human Use (EMA/CHMP). Updated Mar 20 2020. https://www.ema.europa.eu/en/documents/press-release/guidance-sponsors-how-manage-clinical-trials-during-covid-19-pandemic_en.pdf. Accessed Sept 7 2022

7. Points to Consider on Implications of Coronavirus Disease (COVID-19) on Methodological Aspects of Ongoing Clinical Trials. European Medicines Agency Committee for Medicinal Products for Human Use (EMA/CHMP). Updated Jun 26 2020. https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-implications-coronavirus-disease-covid-19-methodological-aspects-ongoing-clinical_en-0.pdf. Accessed Sept 7 2022

8. Guidance on Conduct of Clinical Trials of Medical Products During COVID-19 Public Health Emergency. U.S. Food and Drug Administration (FDA). Updated Aug 30 2021. https://www.fda.gov/media/136238/download. Accessed Sept 7 2022

9. Points to consider on the impact of the war in Ukraine on methodological aspects of ongoing clinical trials. European Medicines Agency Committee for Medicinal Products for Human Use (EMA/CHMP). Updated Apr 13 2022. https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-impact-war-ukraine-methodological-aspects-ongoing-clinical-trials_en.pdf. Accessed Sept 7 2022

10. Fletcher C, Hefting N, Wright M, et al. Marking 2-years of new thinking in clinical trials: the estimand journey. Therap Innov Regul Sci. 2022;56(4):637–50. https://doi.org/10.1007/s43441-022-00402-3.

11. Guizzaro L, Pétavy F, Ristl R, Gallo C. The use of a variable representing compliance improves accuracy of estimation of the effect of treatment allocation regardless of discontinuation in trials with incomplete follow-up. Stat Biopharm Res. 2021;13(1):119–27. https://doi.org/10.1080/19466315.2020.1736141.

12. Polverejan E, Dragalin V. Aligning treatment policy estimands and estimators—a simulation study in Alzheimer's disease. Stat Biopharm Res. 2020;12(2):142–54. https://doi.org/10.1080/19466315.2019.1689845.

13. Lasch F, Guizzaro L, Pétavy F, Gallo C. A simulation study on the estimation of the effect in the hypothetical scenario of no use of symptomatic treatment in trials for disease-modifying agents for Alzheimer's disease. Stat Biopharm Res. 2022. https://doi.org/10.1080/19466315.2022.2055633.

14. Olarte Parra C, Daniel RM, Bartlett JW. Hypothetical estimands in clinical trials: a unification of causal inference and missing data methods. Stat Biopharm Res. 2022. https://doi.org/10.1080/19466315.2022.2081599.

15. Meininger V, Genge A, van den Berg LH, et al. Safety and efficacy of ozanezumab in patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled, phase 2 trial. Lancet Neurol. 2017;16(3):208–16. https://doi.org/10.1016/S1474-4422(16)30399-4.

16. Darken P, Nyberg J, Ballal S, Wright D. The attributable estimand: a new approach to account for intercurrent events. Pharm Stat. 2020;19(5):626–35. https://doi.org/10.1002/pst.2019.

17. Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. Pharm Stat. 2013;12(6):337–47. https://doi.org/10.1002/pst.1549.

18. Little R, Kang S. Intention-to-treat analysis with treatment discontinuation and missing data in clinical trials. Stat Med. 2015;34(16):2381–90. https://doi.org/10.1002/sim.6352.

19. Akacha M, Bretz F, Ruberg S. Estimands in clinical trials—broadening the perspective. Stat Med. 2017;36(1):5–19. https://doi.org/10.1002/sim.7033.

20. Mallinckrodt CH, Bell J, Liu G, et al. Aligning estimators with estimands in clinical trials: putting the ICH E9(R1) guidelines into practice. Ther Innov Regul Sci. 2020;54(2):353–64. https://doi.org/10.1007/s43441-019-00063-9.

21. Mitroiu M, Teerenstra S, Oude Rengerink K, Pétavy F, Roes KCB. Estimation of treatment effects in short-term depression studies. An evaluation based on the ICH E9(R1) estimands framework. *Pharm Stat*. 2022. https://doi.org/10.1002/pst.2214

22. Mehrotra DV, Marceau WR. Survival analysis using a 5-step stratified testing and amalgamation routine (5-STAR) in randomized clinical trials. Stat Med. 2021;40(19):4341–3. https://doi.org/10.1002/sim.9116.

23. Royston P, Parmar MKB. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol. 2013;13(1):152. https://doi.org/10.1186/1471-2288-13-152.

24. Uno H, Claggett B, Tian L, et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol. 2014;32(22):2380–5. https://doi.org/10.1200/JCO.2014.55.2208.

25. Major Depressive Disorder: Developing Drugs for Treatment, Guidance for Industry. U.S. Food and Drug Administration (FDA). Updated June 2018. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/major-depressive-disorder-developing-drugs-treatment. Accessed Mar 7 2023

26. Kessler RC, Bromet EJ. The epidemiology of depression across cultures. Annu Rev Public Health. 2013;34:119–38. https://doi.org/10.1146/annurev-publhealth-031912-114409.

27. Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR). Updated 2022. https://psychiatry.org/psychiatrists/practice/dsm. Accessed Sept 7 2022

28. Medeiros GC, Rush AJ, Jha M, et al. Positive and negative valence systems in major depression have distinct clinical features, response to antidepressants, and relationships with immunomarkers. Depress Anxiety. 2020;37(8):771–83. https://doi.org/10.1002/da.23006.

29. Ten Have M, de Graaf R, van Dorsselaer S, Tuithof M, Kleinjan M, Penninx B. Recurrence and chronicity of major depressive disorder and their risk indicators in a population cohort. Acta Psychiatr Scand. 2018;137(6):503–15. https://doi.org/10.1111/acps.12874.

30. Parker G, Roy K, Hadzi-Pavlovic D, Wilhelm K, Mitchell P. The differential impact of age on the phenomenology of melancholia. Psychol Med. 2001;31(7):1231–6. https://doi.org/10.1017/s0033291701004603.

31. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol. 1967;6(4):278–96. https://doi.org/10.1111/j.2044-8260.1967.tb00530.x.

32. O'Kelly M, Ratitch B. Clinical trials with missing data. New York: Wiley; 2014.

33. Bunouf P, Grouin JM, Molenberghs G. Analysis of an incomplete binary outcome derived from frequently recorded longitudinal continuous data: application to daily pain evaluation. Stat Med. 2012;31(15):1554–71. https://doi.org/10.1002/sim.4491.

34. Estimating Principal Strata. Drug Information Association Scientific Working Group on Estimands and Missing Data. Updated Sept 2 2021. https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data#dia-working-group. Accessed Sept 7 2022

35. Bornkamp B, Rufibach K, Lin J, et al. Principal stratum strategy: potential role in drug development. Pharm Stat. 2021;20(4):737–51. https://doi.org/10.1002/pst.2104.

36. Lipkovich I, Ratitch B, Qu Y, Zhang X, Shan M, Mallinckrodt C. Using principal stratification in analysis of clinical trials. Stat Med. 2022;41(19):3837–77. https://doi.org/10.1002/sim.9439.

37. Lipkovich I, Ratitch B, O'Kelly M. Sensitivity to censored-at-random assumption in the analysis of time-to-event endpoints. Pharm Stat. 2016;15(3):216–29. https://doi.org/10.1002/pst.1738.

38. Boyd AP, Kittelson JM, Gillen DL. Estimation of treatment effect under non-proportional hazards and conditionally independent censoring. Stat Med. 2012;31(28):3504–15. https://doi.org/10.1002/sim.5440.

39. Marwood L, Taylor R, Goldsmith K, et al. Study protocol for a randomised pragmatic trial comparing the clinical and cost effectiveness of lithium and quetiapine augmentation in treatment resistant depression (the LQD study). BMC Psychiatry. 2017;17(1):231. https://doi.org/10.1186/s12888-017-1393-0.

40. Brunner E, Tohen M, Osuntokun O, Landry J, Thase ME. Efficacy and safety of olanzapine/fluoxetine combination vs fluoxetine monotherapy following successful combination therapy of treatment-resistant major depressive disorder. Neuropsychopharmacology. 2014;39(11):2549–59. https://doi.org/10.1038/npp.2014.101.

41. Bell J, Hamilton A, Sailer O, Voss F. The detailed clinical objectives approach to designing clinical trials and choosing estimands. Pharm Stat. 2021;20(6):1112–24. https://doi.org/10.1002/pst.2129.

42. Ratitch B, Bell J, Mallinckrodt C, et al. Choosing estimands in clinical trials: putting the ICH E9(R1) into practice. Ther Innov Regul Sci. 2020;54(2):324–41. https://doi.org/10.1007/s43441-019-00061-x.

43. Ratitch B, Goel N, Mallinckrodt C, et al. Defining efficacy estimands in clinical trials: examples illustrating ICH E9(R1) guidelines. Ther Innov Regul Sci. 2020;54(2):370–84. https://doi.org/10.1007/s43441-019-00065-7.

44. E9(R1) Training Material - PDF_0.pdf. ich.org. Updated Dec 2021. https://database.ich.org/sites/default/files/E9%28R1%29%20Training%20Material%20-%20PDF_0.pdf. Accessed Sept 7 2022