Submission of comments on 'ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials' (EMA/CHMP/ICH/436221/2017)

Comments from:

| PSI/EFSPI/EFPIA |
| --- |
| Contact for clarification: [alan.phillips@iconplc.com](mailto:alan.phillips@iconplc.com)  Co Ordination Team: Alan Phillips, Oliver Keene, David Wright, Rod Junor |

*Please note that these comments and the identity of the sender will be published unless a specific justified objection is received.*

*When completed, this form should be sent to the European Medicines Agency electronically, in Word format (not PDF).*

1. General comments

| Stakeholder number  *(To be completed by the Agency)* | General comment (if any) | Outcome (if applicable)  *(To be completed by the Agency)* |
| --- | --- | --- |
|  | EFSPI/PSI agrees on the importance of clarifying objectives, assumptions and considering how to deal with intercurrent events during the planning and analysis stages of a trial. However, there are several areas where the Addendum can be improved and does not go into enough detail. The impact of introducing the Addendum without clear guidance on how to implement the framework in practice has not been considered and is a serious concern. The main issues to be addressed for a successfully implementation are outlined below:   1. We acknowledge and agree that guidelines need to outline principles. However, there is a need for more case studies to support this addendum. It is proposed that the generic examples in section A.7.A are supplemented by a companion of case studies. The case studies should be a detailed description of a real situation where different strategies for handling intercurrent events are described and the pros and cons of each approach to address them as it relates to the question of interest are articulated. However, it is understood that the drafting group plans to address the lack of examples in the main text by producing case studies in slides that will accompany the Addendum. There are several issues with this approach. Firstly, it is understood that these case studies will be made available shortly, but the final version of the Addendum is unlikely to be completed until the end of 2019. Also, there is no consultation process planned for commenting on the case studies and hence if they could be improved upon there is no opportunity for this to happen, even though there is probably over a year available to do this. It is therefore strongly encouraged that this process is changed to allow comments on the case studies if the Drafting group does not include extensive examples in the main text of the Addendum. 2. The addendum clearly impacts existing disease specific guidance. Regulatory authorities should produce a prioritised list of disease specific guidance documents to revise in light of the Addendum. Given the Addendum will not be finalised until 2019 work should start on revising disease specific guidance documents as a matter of priority to provide better guidance on estimand strategies that may be acceptable in specific disease areas/indications. What plans are there for review of these guidelines once the addendum is issued? 3. It is very important that how intercurrent events have been handled is transparently described in the drug label/SmPC . Please clarify in the Addendum how estimands will link to drug label/SmPC, in particular when a regulatory authority bases its decision on a different Estimand strategy to that pre defined by the Sponsor.   Further, it is recommended the ICH working group provides clarification on the following issues with the Addendum:   1. It is proposed the addendum explicitly highlights that there may be a difference between estimands used for hypothesis testing e.g. ones based on the randomisation with minimum assumptions and estimands used to quantify clinical benefit. 2. The addendum revisits the meaning and role of the analysis populations as outlined in ICH E9; in particular and role of the per-protocol analysis. It is a proposed that a section is added to the addendum that summarises the key changes to ICH E9 and those sections in ICH E9 that no longer apply. That is, clarity can be improved by summarising in one subsection the differences to ICH E9 in the use of ITT and PP analyses as outlined in the addendum. 3. Please clarify whether study discontinuation is an intercurrent event or missing data problem. The Addendum is ambiguous on this issue, and feedback from scientific meetings has not provided clarity. Similarly please can you clarify how death should be handled. 4. The Addendum suggests a Treatment Policy strategy cannot be applied for intercurrent events of death. For a study when there are a small number of deaths “unrelated” to disease/treatment then, please clarify what the appropriate strategy should be? 5. For non-inferiority/equivalence analyses, neither the appropriate strategy for handling intercurrent events nor the appropriate analysis set is specified, and in the case of analysis sets the document seems to contradict ICH E9. Because the inclusion of off-treatment data is likely to bias in favour of no difference between treatments, it seems that the treatment policy strategy and full analysis set are not appropriate in this setting. It is proposed that the addendum explicitly discusses non inferiority trials. A worked example for a non-inferiority trial in the proposed companion of case studies discussed in point 1 above would be helpful. 6. A worked example of an appropriate sensitivity analysis in the proposed companion of case studies discussed in point 1 above would be helpful . 7. Safety estimands: ICH E9 provides guidance specific to the evaluation of safety and tolerability (Section VI). It is proposed this topic is also addressed in the addendum. For example   -are there strategies to handle intercurrent events specifically suited for safety analyses (e.g. ‘while on treatment’)?  -impact of handling of intercurrent events as treatment withdrawals or deaths to determination of number of subjects at risk and estimation of incidence,  -should we analyse safety parameters ‘as treated’ or ‘as assigned’  Moreover, consideration should be given to the possibility that efficacy and safety estimands may use different strategies to handle intercurrent events and this has implications for how benefit risk evaluation is performed.   1. The addendum seems to be based on parallel-group phase III designs. It should be made clear in the introduction that this is the main focus of the Addendum. As well as points 5 and 7 above, some statements should be added to address other settings such as cross over studies, early phase setting, where treatment is given only for a short time like e.g. the conditioning therapy prior to stem cell transplantation in oncology, Bayesian frame work etc . Worked examples for a variety of scenarios in the proposed companion of case studies discussed in point 1 above would be helpful. 2. Existing analytical methods for sample size calculations may not be applicable for some estimation analyses. Therefore, simulation needs to be performed to understand the operating characteristics of various strategies to handle intercurrent events. It is recommended that the Addendum includes a short section highlighting this issue. |  |

1. Specific comments on text

| Line number(s) of the relevant text  *(e.g. Lines 20-23)* | Stakeholder number  *(To be completed by the Agency)* | Comment and rationale; proposed changes  *(If changes to the wording are suggested, they should be highlighted using 'track changes')* | Outcome  *(To be completed by the Agency)* |
| --- | --- | --- | --- |
| 49-50 |  | Comment: This is the only place in this document where the stage of clinical trial is mentioned. Does this imply that the addendum applies to confirmatory clinical trials only?   Proposed change (if any): Clarify whether the scope of the addendum is limited to confirmatory trials. |  |
| 51 |  | Comment: The reference to clinical practice may need more explanation, suggest omit. Also suggest make clear that intercurrent events occur post randomisation.  Proposed change (if any): Randomised trials are expected to be free from baseline confounders, but in practice certain events will occur post randomisation that complicate … |  |
| 51 |  | Comment: The randomisation could also be said to introduce causality by design, by keeping everything, but the treatment, similar in the two groups. If the trial is blinded this will be the case, also during the trial.  Proposed change (if any): Consider to include a sentence consider this causality by randomisation |  |
| 51 |  | “Randomised trials are expected to be free from baseline confounding but…”  Comment: The definition is incorrect.  Proposed change (if any): Change to “Randomised trials are designed to minimize the effect of confounding factors” |  |
| 53 |  | Comment: Here it refers the reader to the Glossary for the definition of ‘intercurrent events’. Actually the Glossary just repeats what is in the previous sentence; albeit using slightly different wording.  Proposed change (if any): Omit reference to Glossary and use glossary wording in the main text. |  |
| 56 |  | “…terminal events such as, in some circumstances, death.”  Comment: Death is always a terminal event, not only in certain circumstances. Also, death seems to be an intercurrent event in all situations as no data can be collected after its occurrence.  Proposed change (if any): Consider to delete “(…) in some circumstances (…)”  Change to “…terminal events such as death.” |  |
| 57 |  | Comment: This is the only instance in the addendum where “safety” is mentioned in the context of variables, or in a broader sense “estimands”. It might be worthwhile to specify whether the considerations in the addendum are meant to apply to efficacy questions only (the addendum is mainly talking about “treatment effect”, which is typically used in connection with efficacy) or whether the considerations equally well apply to safety questions. |  |
| 65 |  | Add ‘i.e.’ before “translating the trial objective into a precise definition of the treatment effect that is to be estimated”, as this is what the term ‘estimand’ describes. |  |
| 70-71 |  | “This addendum clarifies the definition and the role of sensitivity analysis.”  Comment: The addendum should clarify the need for clear specification of the estimand and the criticality and role of sensitivity analyses.  Proposed change: It should be clarified what is the change introduced by the estimand concept with regards to sensitivity analyses. |  |
| 79-81 |  | “However, the question remains whether understanding the effect of a treatment policy always targets the treatment effect of greatest relevance to regulatory and clinical decision making.” Comment: The sentence does not sound clear. Proposed change: Suggest to remove the word “understanding”. Change to “However, the question remains whether the effect of a treatment policy always targets the treatment effect of greatest relevance to regulatory and clinical decision making.” |  |
| 85-89 |  | Lines 85-95 from “On one hand…” repeats material covered in e.g., lines 236ff. |  |
| 87 |  | Proposed Change: “It” should be “they.” |  |
| 91-92 |  | “also between measurements that exist but have not been collected, and measurements that do not, or cannot, exist”.   Comment: The text should more clearly delineate between "cannot" and "do not". "Do not" relates to unmeasured but measurable. "Cannot" is unmeasurable (e.g. post death).  Proposed change: Change to “also between measurements that exist but have not been collected, and measurements that do not exist (ie unmeasured), or cannot exist (ie unmeasurable)” |  |
| 96 |  | Comment: Despite the statement that “Thirdly, the concept of analysis sets is considered in the proposed framework,” the addendum discusses the role of the per protocol analysis set only. Would it be useful to add a small section discussing analysis sets generally? How should an analysis set be defined where, for example, data after initiation of rescue medication are ignored and, in some cases, imputed? Section 5.2 of E9 refers to the analysis set as “the set of subjects whose data are to be included..” but the draft addendum makes clear that, in some cases, not all data on a subject are relevant for every estimand strategy. Should terminology be introduced for this situation – e.g. “modified analysis set” (see Line 840). |  |
| 110 |  | Clarify the robustness context in the document   Proposed Change: Finally, the concept of robustnesss of inferences, …. |  |
| 111-113 |  | Comment: Sensitivity analysis is used in two contexts; is there a better way to differentiate by introducing the terminology of supplementary analysis used later?   Proposed Change: To show how this guidance aligns with supplementary analysis, amend  “In particular, a distinction is made between the sensitivity of inference to the particular assumptions of a particular analysis and the sensitivity to the choice of analytic approach more broadly.” To  “In particular, a distinction is made between the sensitivity of inference to the particular assumptions of a particular estimator and supplementary analyses which investigate sensitivity to the choice of analytic approach more broadly.” |  |
| 120 |  | Comment: What is meant by clear trial objectives? Should there be a one-to-one correspondence between the objective and the estimand? Or could there be several estimands addressing the same objective?  Proposed change (if any): Consider to include more guidance concerning this and/or update figure 1 with more estimands addressing the trial objective if relevant. This could also be included in the example at page 16. |  |
| 122 |  | Comment: Please substitute “handling” with “accounting for” for consistency across the document, since “handling” could be understood in the context of the estimator rather than the estimand. To make the difference between estimand and estimators/ estimates very clear, throughout the whole document, in the context of intercurrent events, the same wording should be used throughout the document.  Proposed change:…the population, the variable, the strategy for intercurrent events… |  |
| 124-125 |  | “The main estimator will be underpinned by certain assumptions.”  Comment: Suggest adding an example for easy reading  Proposed change (if any): Change to “The main estimator will be underpinned by certain assumptions, e.g., no treatment crossover.” |  |
| 128/Figure 1 |  | Comment: should the arrows to intercurrent events originate from “estimand” rather than “estimator”? |  |
| 130-131 |  | … distinguishes between the target of estimation (trial objective, estimand)...: Move trial objectives out of the parenthesis.  Proposed Change: distinguishes between the trial objective, the target of estimation (trial objective, estimand), |  |
| 130-133 |  | “This will assist sponsors in planning trials, regulators in their reviews, and will enhance the interactions between these parties when discussing the suitability of clinical trial designs, and the interpretation of clinical trial results, to support drug licensing.”  Comment: This is already mentioned in Lines 66-68 and 82-83. As stated in general comment please remove the redundancies.  Proposed change (if any): Suggest deleting the sentence. |  |
| 132-134 |  | Add that a clear definition of estimand will allow an upfront understanding on how the data will be analysed and interpreted  Proposed Change: ……when discussing the suitability of clinical trial designs, and the interpretation of clinical trial results, to support drug licensing, prior to the completion of the study. |  |
| 135 |  | Comment: Word missing  Proposed change: In general, it is important to proceed sequentially and not allow for the choice of an estimator to determine the estimand,… |  |
| 135-136 |  | Comment: While this may be true due to the words “in general,” it could be misleading by not acknowledging that there often will be times where consideration of whether there is an estimator that could lead to a reliable estimate may well feed back on the choice of estimands.   This topic was discussed well in Lines 297-308 and 335-338 and the guidance in Line 297 that “an iterative process may be required” is almost in conflict with the guidance in Line 135 that “in general, it is important to proceed sequentially”.  It is recommended that the sentence be followed by a clarification statement.  Proposed change (if any): “In general, it is important to proceed sequentially, and not for the choice of an estimator to determine the estimand, and hence the scientific question that is being addressed. However, there are times when the estimand best reflecting the desired trial objectives cannot be reliably or robustly (i.e. without questionable assumptions) estimated by any feasible design and estimator, and in such cases, alternative estimands that may also address critical, related regulatory questions should be considered.” |  |
| 135 |  | Comment: It is not clear why there is an importance on proceeding sequentially. In practice, the real issue is handing of inter-current events and missing data. Addressing those in the addendum will then make for better estimation.  Proposed change (if any): Consider adding a mechanism to loop back through the framework if estimation issues lead to a re-considering the estimand. |  |
| 141 |  | Comment: The treatment effect described here, as the counterfactual effect of a treatment given compared to when the treatment is denied, to a subject – how does this link to the five strategies described later? For example the treatment policy estimand seems to target an effect of being randomised to treatment -rather than the above described.   Proposed change (if any): Consider to describe how the five strategies can be said to help estimating the described treatment effect or why it is not the aim of the estimand |  |
| 141-144 |  | Comment: Restructure the second sentence (after the colon) for better readability.  Proposed change: A central question for drug development and licensing is to quantify treatment effects. In a specific trial this may come down to the question of how the outcome of treatment compares to what would have happened to the same subjects under different treatment conditions (e.g. had they not received the treatment or had they received a different treatment). |  |
| 144-145 |  | Comment: Intercurrent events do impact a “variable of interest” (which resides in the sample); however, it is more relevant to the discussion to note that these events—or, more precisely, the underlying processes generating the events—relate to the estimand (which resides in the population).  Proposed change: Change “variable” to “estimand” or “population quantity to be estimated.” |  |
| 144 |  | Comment: Please use “definition” instead of “description” for clarification in the context of estimands.  Proposed change: Intercurrent events need to be considered in the description definition of a treatment effect…” |  |
| 151-157 |  | “D. the population-level summary for the variable which provides, as required, a basis for a comparison between treatment conditions”  Comment: Although trial level population is treatment specific free, the treatment level population is specific to, say, a randomized treatment, which is especially relevant for D.  Proposed change: Replace “population“ by “patient set” Change to “D. the Patient-set-level summary for the variable which provides, as required, a basis for a comparison between treatment conditions” |  |
| 151, 159-160 |  | Comment: Clarification is required on how to describe the population based on the inclusion/exclusion criteria i.e. whether this should be the full list of inclusion/exclusion criteria, a cross-reference to the relevant section of the protocol containing inclusion/exclusion criteria or some other appropriate summary. |  |
| 156 |  | Could the wording ‘as estimated by the relevant estimator’ be added here? The estimate and estimator are terms that have been introduced, but it would be helpful to indicate how they relate to the four components of an estimand.   Proposed Change: The population-level summary for the variable (as estimated by the estimator) which provides, as required, a basis for a comparison between treatment conditions. |  |
| 159-162/417 |  | Comment : The phrases “target population” & “study population” might seem to avoid the identifiability problems noted above. However, “target population” is used in another fashion in the EMA draft “Reflection paper on the use of extrapolation in the development of medicines for paediatrics” undergoing public consultation as of Nov 2018. Therefore, continuing to use “target population” & “trial population” (Line 417) in the addendum could lead to confusion eventually, if discussion of any analysis or study requires referring to both the addendum & the reflection paper. Do we need to introduce new terminology for different “sets” of participants in the study so that it is clear who is included in summaries of demography etc and so that “population” is not used in reporting for this purpose? |  |
| 159-160 |  | Comment: If the target population is the patients that are eligible to be included based on the inclusion/exclusion criteria in the protocol, the guideline could clarify if it means that we should exclude from analysis patients randomized but who turned out to have exclusion criteria discovered thereafter. |  |
| 160-162 |  | “In some cases, a stratum of those patients may be of interest, defined in terms of a potential intercurrent event; for example, the stratum of subjects who would adhere to treatment.”  Comment: Restricting the analysis to subjects who adhere to treatment could be in contradiction with the earlier claim that ITT principle should be followed. How to describe the impact on the label?  Proposed change (if any): Clarification should be provided regarding how this scenario would be reflected in the label. |  |
| 160-162 |  | Comment: The use of the word ‘stratum’ is strongly perceived as related to stratified random sampling and this is not the case as described in these lines. Adding the word “subset” might help.  Proposed change (if any): “In some cases, a stratum (subset) of those patients may be of interest ...” |  |
| 163 |  | Comment: Proposed change (if any): Consider adding that these are measurements on an individual (patient) level, as opposed to the “summary measure” attribute which is summarising over patients. |  |
| 167 |  | Comment: Average HbA1c, rather than area under the curve of HbA1c, is a more appropriate measure of clinical benefit.  Proposed change (if any): “for example when using measurements taken prior to discontinuation (e.g., area under the curve of HbA1c until discontinuation average HbA1c over time until discontinuation;” |  |
| 172-173 |  | “For example, if a subject dies before a planned measurement of blood pressure, the blood pressure will not be observed”.   Comment: This is a scenario of ‘cannot’ and should be unambiguous.  Proposed change: Change to ““For example, if a subject dies before a planned measurement of blood pressure, the blood pressure cannot be observed”. |  |
| 175-176 |  | Comment: The last part is not true in general. Please substitute “will” with “might”.  Proposed change: If a subject discontinues treatment because of toxicity, the blood pressure may be observed but might reflect the lack of effect of the treatment when it is not taken. |  |
| 178-180 |  | Comment: Please substitute the first “and” with “with” for better readability (due to the second “and” in the sentence).  Proposed change: Taking use of rescue medication as an example, two different specifications include the combined effect of treatment and with any intercurrent event… |  |
| 182 |  | Comment: The intercurrent events such as discontinuation of treatment due to lack of efficacy or AE; or introduction of rescue medication, may reflect the trial design rather than clinical practice. If the estimated treatment effect is depends heavily on the strategy for dealing with intercurrent effects – will it then be relevant for a future patient, who will have a different risk to experience similar intercurrent events as observed in the trial?  Proposed change (if any): Consider to include a discussion of the interdependence between trial design and occurrence of intercurrent events |  |
| 183-187 |  | Comment: This paragraph discusses the summary measure for the variable, the fourth element of the estimand full description. Neither this section nor other sections clearly describe the distinction between the summary measure component of the estimand, and the estimator depicted in Figure 1 (Line 128). It is recommended that the distinction between these is clarified using a generic example.  Proposed change (if any): “…under two different treatment conditions. The estimator (Figure 1) is distinct from the population level summary measure; the estimator is a specific statistical method for calculating the estimand. For example, the summary measure of an estimand may be stated as the mean change from baseline (of variable X at time T), and the estimator might be ‘…. Calculated with ANCOVA using covariates of A, B, C’ .”] |  |
| 187 |  | …..under two different treatment conditions.  Proposed Change: …..under two different treatment regimens. |  |
| 188 |  | In the section on “Strategies for addressing intercurrent events”, it will be helpful if the document provides guidelines on how to handle missing data that was resulted from the intercurrent event (for example, the subject took rescue medication but it was not able to collect data afterwards). |  |
| 189 |  | Depending on whether this document should follow US or UK English, replace ‘through’ with ‘to’. |  |
| 189-190 |  | “The estimand attributes A through D introduced in Section A.3.1 are inter-related and should not be considered independently.”  Comment: A sort of introduction to intercurrent events is provided in paragraph starting line 40. However, no clear definition is provided.  Proposed change: Add clear definition of intercurrent events here. |  |
| 189 |  | The term inter-related is not clear: it should be removed and completed by the chart with the four circles defining the estimands attributes for more clarity.  “The estimand attributes A through D introduced in Section A.3.1 should not be considered independently”. |  |
| 190-191 |  | Comment: Please substitute “reflecting” with “accounting for” for consistency Throughout the document, in the context of intercurrent events, for the same reason outlined in the comment for line 122  Proposed change: The description of an estimand will not be complete without accounting for how… |  |
| 197 |  | Comment: Minor alteration of the sentence to include another important aspect to consider.  Proposed change: The relevance of each strategy will depend on the objective as well as on the therapeutic and experimental context. |  |
| 198 |  | What is an ‘experimental situation’ in this context - is it referring to the type of clinical trial? |  |
| 200 |  | Section A3.4 is noted. That section does not exist. |  |
| 201 |  | Wouldn’t an adequate description of the chosen strategy to be used be the estimand itself? Are we therefore allowed to deviate from the 5 descriptions given in the document? Could this get confusing when trying to compare different studies? |  |
| 204-212 |  | Comment: Use of the words “irrelevant” and “ignored” are appropriate if using an ITT approach to estimating a treatment policy. If a marginal structural model is employed, these intercurrent events are accounted for as part of the estimand.  Proposed change (if any): "The value of the variable of interest is used regardless of whether the intercurrent event occurs. For example, when specifying how to account for rescue medication as an intercurrent event, observations on the variable of interest are used regardless of whether rescue therapy was taken. |  |
| 208-209 |  | Comment: This document, appropriately, does not name any specific strategy as leading to what should be named the “intention to treat analysis” (ITT), and only notes that the treatment policy strategy ‘reflects’ the description in ICH E9 for ‘intention to treat principle’. There remains in many venues unclear usage of the term “ITT”, and often regarded that the treatment policy strategy for intercurrent events is the only approach that can lead to a “ITT analysis”. This sentence in the guideline may seem to imply this as well. This guideline will add valuable clarity to study design and analysis protocols and discussions by improving clear understanding of the term.   Proposed change: “If applied across all types of intercurrent events, this reflects the comparison described in the ICH E9 Glossary (under Intention to Treat Principle) as the effect of a treatment policy. The term ‘ITT principle’ should be used to mean the principle of including all randomised study subjects in the analysis set, using data and methods as defined by the estimand. Furthermore, there is no single strategy for intercurrent events that should always be applied for an analysis to be consistent with the ITT principle, nor should any specific analysis be named as ‘the ITT analysis’. “ |  |
| 210-211 |  | Comment: Clarify whether “do not exist” means that it’s not possible for the event to exist (i.e. following death) or missing data for other reasons (e.g. lost to follow up as subject left the country and didn’t want to continue). Is death the only reason for measurements not to exist?   As written this section implies the treatment policy strategy cannot be applied to a study where even a single subject dies. If deaths are unrelated to treatment/disease this seems unreasonable. Please clarify if this is the intent.  Proposed change (if any): In general this strategy cannot be implemented when values for the intercurrent event do not exist for all subjects (unless the number of subjects is sufficiently small that this can be disregarded). |  |
| 210-212 |  | Comment: It is stated that an estimand based on the treatment policy cannot be constructed with respect to a variable that cannot be measured due to death. However, even in a trial of a non-life-threatening disease, it cannot be excluded that death may occur, in particular in trials with large sample size and/or long duration. Does that mean that this strategy can never be fully achieved, and that other strategies, such as composite strategy or hypothetical strategy (eg. had no patient died) should be considered for such trials? |  |
| 210-212 |  | …this strategy cannot be implemented when values for the variable after the intercurrent event do not exist for all subjects.  Proposed Change: Could we also state that ‘In the case of values that were planned to be collected but were not, suitable imputation methods can be used (see section A7.1)’ |  |
| 217 |  | Comment: both “multiple” and “different” are used to describe approaches  Proposed change (if any): ”There are multiple approaches that can be considered under this label.” |  |
| 217-224 |  | Comment: An example of a composite strategy would help the reader to better understand the concept. We propose inclusion of an example such as the one below in this section.   Proposed change: A composite strategy may be implemented in HCV trials where data on response are collected after treatment discontinuation? If a subject prematurely discontinues treatment but still responds at PT Week 12 he is considered a responder. On the other hand, response data after rescue medication initiation are considered non-responders. |  |
| 221-223 |  | The terminology “numerical variable” should be clarified. Does it mean ordinal? Continuous or both? i.e. is it allowed to use a composite strategy with a continuous endpoint (provided you can define what a “extreme unfavourable value” is for the continuous endpoint)? |  |
| 223-224 |  | Comment: The use of an AUC based on values prior to the intercurrent event does not seem to reflect the composite strategy. It is for example mentioned as an example of the while on treatment strategy in line 743). This strategy seems also biased in case the variable would spontaneously deteriorate over time, which would be the case in progressive diseases (such as Alzheimer, Parkinson…) |  |
| 232 |  | Comment: Under what regulatory setting will the hypothetical strategy be appropriate?   Proposed change (if any): Give specific situation(s) under which a regulator might be interested in the hypothetical strategy. |  |
| 233 |  | “A scenario is envisaged in which the intercurrent event would not occur”  Proposed change: “A hypothetical scenario is envisaged with regard to the intercurrent event” |  |
| 233-235 |  | Comment: Suggested change in text for clarity.  Proposed change (if any): “… the value to reflect that scientific question of interest is that the value which the variable would have taken in the hypothetical scenario defined.” |  |
| 236-237 |  | Comment: Please add a clarification what “not been available” means as we think this could be any of the items below.  Intercurrent event never occurred or Rescue medication never approved Rescue medication not been made available to the subjects |  |
| 244-245 |  | Comment: As above, an example will help the reader understand the concept here. We propose inclusion of an example such as the one below in this section.  Proposed change: For example, the data after rescue medication was initiated would be excluded, and the model and/or imputation method used should assume or impute missing data as if the results were continuing in the same trend as before the rescue medication was initiated. |  |
| 245-247 |  | Comment: Please correct important typo which alters the sentence meaning.  Proposed change: For example, the hypothetical condition might usefully address both the use of a rescue medication and *non-adherence* to treatment as intercurrent events in order for an estimand to be precisely described. |  |
| 247 |  | Could we also add that suitable imputation methods may be used to reflect the hypothetical condition (see section A7.1). |  |
| 248 |  | Comment: The guidance for using the principal strata strategy in the current version is very limited.   Proposed change (if any): In the example section suggest to add suggestions for what steps would be involved in doing such an analysis |  |
| 248-263 |  | Comment: It looks like that unless eligible patients are identified before randomization, for example based on a run-in period (i.e. a randomized withdrawal trial), this strategy will always rely on untestable assumptions and that no robust estimator can be proposed (see also later comments on lines 728-735). This makes the relevance of this estimand questionable. It is also unclear what would be the impact on the labelling on this estimand. |  |
| 248-263 |  | Comment: It will aid clarity to note the distinction between the other strategies which address what observed data to include in the analysis set and how to address missing data to enable use of the analysis set data, versus the principle stratum strategy which addresses intercurrent events by defining which study subjects should be included in, or entirely excluded from, the analysis set.   Proposed change: “… because different subjects will experience different intercurrent events on different treatments. The principle stratum strategy differs from several other strategies by defining which study subjects will be represented in the analysis set, rather than how to include the observed data or occurrence of missing data in the analysis set.” |  |
| 249 |  | Comment: Line 249 introduces the principal stratum approach, and states that the principal stratum can define the target population. This is inappropriate. As a fundamental principal, the target population should be one that a treating physician can identify; however, the principal stratum cannot be identified by a physician.  There is a subsequent example on line 423, where the principal stratum consists of patients who tolerate the treatment. Since patients must be treated in order to determine who tolerates it, it is of questionable use to a physician that the group of patients who should receive the medication is the group that tolerates it.  It would not be sufficient to claim that physicians should simply treat all potential patients until tolerability is determined, because the effect (positive or negative) of the treatment in the subset of patients who don’t tolerate it would need to be accounted for. |  |
| 251-3 |  | Comment: As written, the sentence “In other words, a principal stratum is a subset of the broader population who would not experience the intercurrent event” is misleading since it implies that a principal stratum is always a subset of the population in which an intercurrent event would not occur. With reference to the previous sentence (line 250-1), the principal stratum of interest is the subset of the population who would not experience the intercurrent event on either treatment.   Proposed change (if any): “In other words, the relevant principal stratum in this case is the subset of the broader population who, would not experience the intercurrent event” |  |
| 254 |  | Please better clarify the difference between principal strata and subgroups. |  |
| 255-261 |  | Lines 255-261 from “Principal stratification….In contrast” repeats the definition in the Glossary. |  |
| 255, 714, 857 |  | Comment: "... patient's potential intercurrent events on both treatments ..." The Glossary defines four possible principal strata. However, an implementation of this would be the patients who would not have had an intercurrent event of treatment withdrawal on treatment B (the novel treatment) irrespective of whether they had an intercurrent event on treatment A (the standard of care). An example of which would be a treatment for a chronic condition for an estimand of interest to a Payer who will only be interested in those patients who are taking the novel treatment long-term in comparison to a treatment policy estimate of the efficacy on the standard of care in this population. This seems a more useful analysis than the no intercurrent event under either treatment which is primarily discussed in the amendment, which is a harder population to conceptualise.   Proposed change (if any): Consider adding the above as an example of using the principal strata strategy. |  |
| 259 |  | Comment: "... randomised controlled trial because each patient will be observed on one treatment only". This is true for a parallel group study. However, in a cross-over study these patients could be identified (under some assumptions at least, e.g. on wash-out). |  |
| 260 |  | Please clarify what is meant by “inferred from covariates”. How is this done? |  |
| 260 |  | Example(s) of the statistical model that should be used in that case should be provided. In addition the use of wording "as imperfectly" seems to indicate that the approach is controversial, in that case is it pertinent to mention it? |  |
| 264 |  | The name while on treatment strategy could suggest that "last value under treatment" could fall under this heading. Consider to make a more explicit statement whether or not this approach could be considered as a "while on treatment strategy". |  |
| 265-267 |  | “If a variable is measured repeatedly, its values up to the time of the intercurrent event may be considered to account for the intercurrent event, rather than the value at the same fixed timepoint for all subjects.” Please clarify, are the measurements up to the intercurrent event considered to account for the intercurrent event or rather to account for the response to treatment? If no imputation of measurements after the intercurrent event is needed then please explicitly clarify. |  |
| 265 |  | Comment: For clarity, please consider minor alteration to the sentence.  Proposed change: Only response to treatment prior to the occurrence of the intercurrent event is of interest. |  |
| 265-267 |  | Comment: The content of this sentence relating to how to deal with different durations of observations for patients or treatments requires further detail as the meaning was not clear to us as a reader. |  |
| 266-267 |  | Comment: The wording/sentence structure is ambiguous.  Proposed change (if any): “… its values up to the time of the intercurrent event may be considered to account for the intercurrent event the only relevant values in this strategy rather than the value at the same fixed timepoint for all subjects.” |  |
| 272-276 |  | Comment: This paragraph is not a continued description of the topic ‘While on treatment strategy’, but a final overview of the all topics under Section A.3.2. It is recommended that this paragraph becomes a separate summary, preceded by a heading. The other comments provided here relating to this overview/summary will also be more clearly relevant if this is added.   Proposed change: “Diversity and precision of strategies   Altogether, five different strategies are considered in this section.” |  |
| 272 |  | Comment: The text should not be restricted to 5 possible strategies, the possibility of adding new strategies that might be devised in future should be allowed for.  Proposed change: “Altogether, five different strategies are considered in this section. Other strategies for an intercurrent event are not precluded, and should be considered if relevant and appropriate.” |  |
| 272-276 |  | Comment: Section A7 provides simple generic examples of applying strategies to create an estimand. Even the examples with two intercurrent events are simplified to enable concise presentation. Actual studies, however, will often have multiple types of intercurrent events, each of which needs to have a strategy selected to create a comprehensive estimand. This paragraph, or new paragraph within the newly named subsection, should state the expectation that real study planning will often require explicit identification of all types of expectable intercurrent events, and a diversity of strategies across the intercurrent events might be selected to address the events. Although this thought is briefly expressed in a later portion of the document (Line 288), this will be a valuable location to also express it to avoid potentially leaving the impression that the just described strategies are to be selected among for uniformly applying to all events for any individual estimand.  Proposed change: “… (iii) the effect during adherence.  Actual studies, particularly long or complex, or in complex clinical circumstances, may be expected to have multiple types of intercurrent events. A well-defined estimand will identify a strategy for each type of intercurrent event. Often a study objective will be best served by employing a variety of strategies across the event types in defining an estimand.” |  |
| 279-285 |  | Lines 279-285 to “…trial results” repeat material that is covered elsewhere, e.g., in lines 144-150. |  |
| 281-283 |  | “The construction of the estimand should address each intercurrent event that may occur in the clinical trial and that will affect the interpretation of the results of the trial”   Comment: That sentence is in conflict with Lines 285-286 (“It may be impractical to foresee every relevant kind of intercurrent event”).  Proposed change: Change to ““The construction of the estimand should address the practically foreseeable intercurrent events that may occur in the clinical trial and that will affect the interpretation of the results of the trial” |  |
| Lines 281-283 |  | Comment: While this makes sense, in practice fine-grain differentiation will be statistically extremely difficult or impossible and so multiple intercurrent event types will need to be grouped together: Most clinical trials will have insufficient patients to allow for differential handling of more than 2-3 types of intercurrent event in a statistically sound manner (e.g. by within-group imputation). It can also be extremely difficult to objectively classify many types of these events; stated reasons may not be the real underlying reason (for instance, patient withdrawing consent). A patient might be labelled as discontinued, but then may or may not have gone on to take other medication. Although these issues may be reduced by improved training and recording, they are difficult to remove entirely as classification of many types of intercurrent events is subjective and/or ambiguous. In general, even when handled sensibly, analysing different types of intercurrent event in different ways will result in increased statistical complexity, variance and likelihood of analysis failure  Proposed change (if any): We would welcome guidance on this topic, possibly including a framework/structure for recording and classifying intercurrent event types in an objective manner, and potentially standard groupings. |  |
| 285 |  | Comment: Please add further clarification why these specific criteria do not affect interpretation of trial results, since we cannot see why such criteria are not expected to affect the interpretation of trial results. |  |
| *286-288* |  | *Comment: “the effect on what the chosen analysis estimates” is quite unclear. An example might be useful.* |  |
| 297 |  | Comment: Suggested change in text for clarity.  Proposed change (if any): “An iterative process may be required to construct the estimands of interest.” |  |
| 299-301 |  | Comment: Minor alteration of the sentence for clarity.  Proposed change: “…which is reliable for inference regarding the estimand can be derived.” |  |
| 300 |  | Comment: Methods of data collection (including accurate recording of the occurrence of intercurrent events) must also be considered.  Proposed change (if any): Suggest add “and methods of data collection” |  |
| 302-305 |  | “Some estimands, in particular those that are estimated using the observed data, can be robustly estimated making few assumptions, whereas other estimands require more specific assumptions that may be more difficult to justify and that may be more sensitive to plausible changes in those assumptions”  Comment: It seems the clause,“…in particular those that are estimated using the observed data…” is confusing since any estimand would use observed data. It is just that some estimands may also use external data (e.g. those incorporating the hypothetical or principal stratum approaches).  Proposed change: Change to “Some estimands, in particular those that are estimated using only observed data from the study, can be robustly estimated making few assumptions, whereas …” |  |
| 302-308 |  | Comment: There is much emphasis on the robustness of estimating using the observed data. This might not be true in some situations. For example, in rheumatoid arthritis, subjects who do not respond well are allowed to escape to the test treatment. If a treatment policy is requested for this “escape” intercurrent event, the comparison of interest might become between the drug and same drug with a delayed start. A hypothetical strategy might be more useful.   Proposed change (if any): “… trial design and analytic approach would need to be considered. In some circumstances the estimand most ‘robustly estimated’ may be not addressing a useful question and a less robust, but more relevant estimand may be preferred as the primary estimand.” |  |
| 304 |  | Comment: Clarify which estimands are considered able to be robustly estimated, and which may be more sensitive to changes in assumptions. Are treatment policy and composite estimands the robust ones and hypothetical, principal stratum the less robust ones? |  |
| 306-308 |  | In the described situation, is a discussion in the study protocol expected that another estimand would be more appropriate, but cannot be reliably estimated? |  |
| 306 |  | Where significant issues exist to develop an appropriate trial design or to derive a reliable estimate for a particular estimand, an alternative estimand, trial design and analytic approach would need to be considered.  It would be appreciated to clarify in an example when different estimands might be possible to evaluate different strategies for addressing intercurrent events. E.g. for Oncology would it be possible to use overall survival time as estimand for treatment policy strategy (de facto estimand) and PFS time as estimand for hypothetical strategy (de jure estimand)? This would avoid the need of making several assumptions difficult to check when using methods like Rank Preserving Structural Failure Time |  |
| 318-326 |  | Comment: This discussion begs the question of a “robust estimate” ever exists in cases where we need to model what would have happened to a given subject under a different treatment.   Proposed change (if any): Provide one or more specific examples of robust estimators in these more challenging settings. |  |
| 329-331 |  | Comment: We modified the sentence because we did not understand why “for use in treatment naïve subjects” is mentioned and thought it was possibly a mistake as the sentence has the same meaning for non-treatment naïve patients.  Proposed change: If the treatment is proposed for use in treatment naïve subjects as part of a treatment policy… |  |
| 331-334 |  | Comment: In some clinical trials, the number of subjects with an intercurrent event could be small. Statistical inference suggested by the addendum in Lines 331-332 for an additional estimand and analysis pertaining to the intercurrent event could be misleading due to small sample size and lack of the power.  Proposed change: … inference can be complemented by defining an additional estimand and summary analysis pertaining to the intercurrent event itself… |  |
| 335-338 |  | Comment: The recommendation to use treatment policy estimands even when they are known to be less clinically-relevant is strongly disagreed with, and would result in worse regulatory decisions. If the estimand of clinical interest were hypothetical, then use of treatment policy as a surrogate estimand would not improve decision making. This is because the treatment policy estimate will typically be used and reported without allowance made for it not reflecting the most clinically relevant estimand, and the significance testing upon which regulatory decision making is primarily based cannot account for this discrepancy in ‘meaning’. Use of treatment policy in this case therefore does not remove bias, it merely moves it, hides it, and amplifies it: The known statistical biases of hypothetical estimation (whose potential impact is quantifiable by sensitivity analysis) are exchanged for clinical bias in treatment policy, which is harder to identify or quantify as it occurs outside a statistical framework, but which logically must on average be greater in magnitude (else the treatment policy estimator would also be a less biased estimator of the hypothetical estimand than the hypothetical estimator itself).  Where treatment policy is less clinically relevant, the use of it due to its ‘unbiased estimation’ in practice simply represents the introduction of a strong conservative bias to the estimation of the desired (e.g. hypothetical) estimand. Where the primary concern is robustness and there are no better methods available, the correct response is to perform sufficient sensitivity analysis to assess the impact of reasonable deviations of assumptions. As this is already mandatory good practice, there is no need to deliberately choose clinically less-relevant estimands.  Proposed change (if any): Remove this recommendation entirely. |  |
| 335-342 |  | Comment: This paragraph states that in some cases the estimand based on the treatment policy strategy may be the one best suited to support regulatory decision making (e.g., hypothesis testing) because it may best support robust inference when estimands of greater clinical interest cannot be formulated with similar robustness. The sentence encouraging use of a treatment policy estimand is then stated to be “still relevant”. This phrasing may be misunderstood to mean that if the treatment policy estimand supports robust inference and regulatory decision making, then it is also the relevant estimate to consider for all purposes. The last sentence in the paragraph only partly ameliorates that.   Proposed change: “… that are agreed to support a reliable estimate or for robust inference. An estimand based on the treatment policy strategy might offer the possibility to obtain a reliable estimate of a treatment effect that is still adequately relevant. In this situation, it is recommended to retain those estimands that are considered to be of greater clinical relevance and to present the resulting estimates all estimands and resulting estimates along with a discussion of the relevance and limitations of each, in terms of trial design or statistical analysis, for that specific approach. |  |
| 336-338 |  | Comment: It is unclear what is meant by “specifically in settings where estimands based on alternative strategies might be considered of greater clinical interest, but main and sensitivity estimators cannot be identified that are agreed to support a reliable estimate or robust inference". Does it refer to “Principal stratum strategy”. Does it mean that treatment policy strategy would be used instead of this strategy? Are there any case studies that underlie this statement in the draft ICH E9 (R1)? |  |
| 341-342 |  | …present the resulting estimates along with a discussion of the limitations..’  Would it be possible to indicate that this information is also useful to include in a label, particularly for physicians/ |  |
| 343-348 |  | Comment: Please do not recommend dichotomizing a continuous variable. There is substantial literature discussing the fact that this approach has tremendous cost in statistical efficiency, while failing in its goal to address the question of clinical meaningfulness.  Proposed change (if any): Balance this section with the disadvantages of dichotomizing a continuous variable and include the possibility of retaining discrimination in the continuous variable |  |
| 370-371 |  | Comment: These two lines seem not to correctly connect. Likely, there is a full stop at the end of line 370 that should not be there.  Proposed change (if any): Remove full stop at the end of line 370 or re-word:  Use of a treatment prohibited by the protocol or use of a subsequent line of therapy will commonly … |  |
| 377-378 |  | “The choice of estimands for studies with objectives to demonstrate non-inferiority or equivalence requires careful reflection”  Comment: This sentence adds little to the exposition. All elements of study design require careful reflection.  Propose change: Suggest deleting this sentence. |  |
| 377-390 |  | Comment: While the rest of the document stays away from specifying definitions of analysis sets, this concept is introduced in this paragraph with a discussion around the FAS and no mentioning of the Per Protocol Analysis Set. It is unclear what strategies could be employed for the intercurrent events of “protocol violations and deviations, non-adherence and withdrawals “ i.e. whether subjects with these events should be eliminated from the population definition in one type of estimand and, if so, how this type of estimand might be defined. |  |
| 377-390 |  | Comment: This paragraph discusses considerations for non-inferiority trials. Another possible consideration concerns determination of the non-inferiority margin, which is often based on a historical trial comparing the active control to a placebo.   Proposed change (if any): Clarify or discuss what estimand should be used to estimate the active control’s effect. There may be some rationale for using the same estimand in the historical trial and the non-inferiority trial. |  |
| 378-381 |  | Comment: For clarity, it is recommended that the term “conservative” be defined. |  |
| 379 |  | Proposed change: ‘he’ should be ‘the’ |  |
| 388-390 |  | Comment: It would be good to have an example of “a measure of treatment effect with high sensitivity to detect differences between treatments.” More generally, more specific advice on analysis of non-inferiority studies would be helpful. |  |
| 391 |  | Comment: is the title (“Impact on trial ‘sign’ and conduct”) correct?  Proposed change: Should ‘sign’ be ‘design’? |  |
| 393 |  | Comment: Would it only be the primary trial objectives? As stated later in the addendum it is important that the study is designed appropriately for the secondary (including safety) objectives too.  Please provide clarity on what is intended by primary objectives – We interpret this to mean primary and key secondary endpoints for which a labelling claim will be made – is this correct?  Proposed change (if any): Please clarify that this applies to key secondary endpoints too. |  |
| 408 |  | “untestable assumptions and, depending on the proportion of missing data; this may undermine the…”  Comment: There is a typo (comma, not colon)  Proposed change (if any): change to “and, depending on the proportion of missing data, this” |  |
| 409 |  | Comment: It seems not easy to distinguish “intercurrent events” and “residual missing data”.  Proposed change (if any): Consider adding more examples of what would be considered “residual missing data”. |  |
| 413 |  | Comment: If “treatment discontinuation due to lack of efficacy” is considered an intercurrent event of interest, it would generally be helpful if the criteria that constitute lack of efficacy were determined in advance.  Proposed change (if any): Change sentence to “For example, perhaps a generic “loss to follow up” should correctly be recorded as “treatment discontinuation due to lack of efficacy”, with lack of efficacy criteria defined in the protocol.” |  |
| 413 |  | Comment: The use of the term ‘correctly’ in this context is not appropriate, the term ‘accurate’ is clearer and more correct.   Proposed change: …perhaps a generic ‘loss to follow up’ should be more accurately be recorded as ‘treatment discontinuation due to lack of efficacy’. |  |
| 415-419 |  | Comment: In general, a clinical trial in itself is not reflecting clinical practice. There are in general more visits and more assessments/interventions in a clinical trial compared to normal clinical practice and this in itself is likely to impact retention of subjects. Furthermore, we conduct multi-regional trials and clinical practice differs across regions and countries. We need to adhere to the highest standards in the participating countries. Also, when comparing e.g. two insulins, we need titration targets/schemes in order not to favour one insulin to the other and this titration needs to be monitored closely and action taken if no good explanation exists for a deviation from the titration algorithm. This type of titration is typically not reflecting clinical practice, but something we need to implement to ensure sufficient titration and avoid biased comparison.  Proposed change (if any): Please consider softening the first sentence and delete “titration schemes” from the example. |  |
| 420-429 |  | Comment: Sequential multiple assignment randomized trial (SMART) designs in which patients requiring rescue therapy are re-randomized to a specific rescue therapy, allow for the estimation of treatment policy estimands for a specific frontline and a specific rescue therapy.  Proposed change (if any): Add SMART designs to the list of examples of non-standard trials in row 421. |  |
| 421 |  | Consider adding cross-over designs (completers) |  |
| 424 |  | It seems that a randomized withdrawal design is described here. Suggest to write "(..) subjects who can tolerate a treatment using a randomized withdrawal design with a run-in period (..)". |  |
| 425-426 |  | The dialogue should also agree on the treatment to be used in the run-in period. More in particular, whether it could be placebo. |  |
| 430-434 |  | Comment: The impact on sample sizes could be discussed more deeply. In many situations it could be difficult to target an effect-size that accounts for the impact and handling of intercurrent events. The impact on sample size and possible inflation should be discussed in such situations. An example would be of interest. |  |
| 435 |  | Change to Section A.7.2. But also check, is that reference accurate? Did you mean Section A.7.A instead? |  |
| 435-443 |  | Comment: This paragraph mentions the need to have consistent definition for the variable of interest, but does not mention the potential impact of different intercurrent events across trials. |  |
| 444 |  | “More generally, a trial is likely to have multiple objectives translated into multiple estimands.”  Comment: Sample size should account for intercurrent events. Therefore, it is not clear how to include intercurrent events in sample size estimation.  Proposed change: Consideration and guidance on how to address multiplicity issue and impact on sample size should be provided. |  |
| 444-446 |  | Comment: It would be appropriate to acknowledge that the treatment policy estimand is not always required.  Proposed change (if any): “A trial design that is suitable for one estimand might not be suitable for other estimands of potential importance. In addition, the treatment policy estimand is not always required. Trials with multiple objectives and endpoints …” |  |
| 454 |  | Comment: Clarification is suggested.  Proposed change: “… so that the effect of treatment can be isolated from any treatment unrelated differences between the groups of subjects on which the comparison is based.” |  |
| Lines 462-471 |  | Comment: The primary concerns around the robustness of estimation of hypothetical estimands (lines 468-471) arise from handling of missing data; without it, statistical assumptions are essentially the same as with treatment policy. However, in real data, even with the best plans possible for patient follow-up after discontinuation/rescue etc., there will always be missing data in later phase clinical trials. Lines 462-464 therefore present an unduly positive and misleading picture of treatment policy (in the presence of missing data, the complexity of analysis and assumptions are arguably greater for treatment policy). Lines 465-468 state that a composite endpoint may need no further statistical assumptions regarding missingness, even when it is present. However, this document encourages the differentiation between missingness and intercurrent events (e.g. lines 483-485), and it is likely that some of the missing data cannot be treated as ‘failure’ in a clinically reasonable way (e.g. where no intercurrent event had occurred). Therefore, treatment policy and composite strategies may reduce missing data, but in practice are unlikely to lead to its elimination. Wherever there is missing data, assumptions automatically become strong and untestable, and hence all five estimands strategies outlined require the same types of assumption that are so criticised in this document. That these assumptions have to be made is therefore inevitable and should not be criticised as a flaw of any strategy. Where the strategies differ is in the amount of (relevant) missingness that they generate, and therefore the sensitivity and robustness of their analysis to the assumptions. This is a classic missing data (MNAR) issue which should be handled (for all primary analyses) by sufficient sensitivity analysis.  Proposed change (if any): Cover the problem of missing data requiring strong, untestable, assumptions without reference to specific estimands (since it applies to them all). Focus on use of sensitivity analyses to assess robustness of results to these assumptions. It would be fair to state that some estimands are more robust to deviations in these assumptions than others. |  |
| 467-468 |  | Comment: Not a logical sentence, edited for clarity.    Proposed change: Using a composite strategy it may be possible to perform an analysis without need for imputation or modelling of response after an intercurrent event and without the associated assumptions of such modelling or imputation.” |  |
| 468-469 |  | Comment: Did the author intentionally distinguish between "a strategy that requires a hypothetical scenario" and "the hypothetical strategy"? |  |
| 471 |  | Comment: The sentence starting “In a randomized….”. Unclear what this is telling me and not sure if it’s correct in all cases.  Proposed change (if any): Rephrase or remove this sentence. |  |
| 472-473 |  | “In a randomised trial, estimation of a treatment effect within a principal stratum of the population will be confounded unless the subjects within that stratum can be identified before randomisation.”  Comment: It is unclear what the sentence is supposed to illustrate. |  |
| 474-476 |  | Comment: A comparison of adherers on drug vs. control could be informative. Regardless of the reasons for non-adherence a notable difference in adherence to therapy is an important observation in considering the utility of the investigational drug especially in a study with an active comparator. The fact that they are different and what makes them different may also be informative for treatment decisions should the investigational drug eventually be authorised for use.   Proposed change (if any): Clarify that it is only inappropriate to compare outcomes in these two strata. |  |
| 477-479 |  | Comment: These lines seem to suggest that a “preconceived” set of covariates is undesirable, but pre-specification of model terms is generally regarded as good statistical practice.  Proposed change (if any): Clarify whether the covariate list should be pre-specified or not. |  |
| 479 |  | Comment: Clarification regarding the intended meaning of the word ‘labelled’ is requested. |  |
| 480 |  | Comment: Suggest spelling out what you mean by “stronger assumptions”. |  |
| 491-496 |  | Comment: Prediction model based on other patients who discontinued treatment but for whom data collection continued is probably a good approach, but requires that a sufficient number of patients are available to build a stable model. If we assume for example a 10% discontinuation rate, and 50% of these patients have data available, this would lead to build a model from 5% of patients, which will likely be unstable and result in highly variable predictions. Would control-based imputations or imputations based on external data be considered an acceptable approach, providing the underlying assumptions are justifiable? |  |
| 500-501 |  | Comment: This statement may be misinterpreted to mean that estimates used for inference (e.g., hypothesis testing) should be ‘absolutely’ robust. The need, often, to balance reliability and relevance may lead to using an estimate with high relevance but not the highest reliability as the best available option. Qualification of the term ‘robust’ is thus recommended.  Proposed change: “Inferences based on a particular estimand should be adequately robust to limitations in the data and deviations from the assumptions used in the statistical model for the main estimator. |  |
| 508-509 |  | "Each supplementary analysis may refer to a different estimand or a different estimator of the same estimand." The last part is confusing, as estimators focusing on the same (i.e., main) estimand can better labelled as "sensitivity estimators". |  |
| 509 |  | It should be useful to add an example of different estimators to the same estimand for a supplementary analysis. In our understanding the use of different estimators for the same estimand should be considered only as sensitivity analyses so maybe the sentence should be reworded. |  |
| 537-559 |  | Comment: It is unclear what is the difference between supplementary analysis (eg. targeting a different estimand from the same variable or endpoint) and estimand for secondary trial objectives (eg. for a secondary variable or endpoint) Could it be clarified whether it is expected that results from supplementary analyses should confirm the conclusions from the primary analysis (to match what was done before with FAS and PP set), or that since it is addressing a different question, different results may be expected. |  |
| 548-550 |  | Comment: This section states that consistent results between the full analysis set and per protocol set increase confidence in the study results, but also states that the per-protocol results can have “severe bias”. It is not clear why consistency with a potentially severely biased estimate would be reassuring, or conversely, why lack of such consistency would make one less confident about a study’s results.  Proposed change (if any): Resolve this apparent contradiction. |  |
| 556-559 |  | Comment: The use of a per-protocol analysis, especially in the context of non-inferiority/equivalence trials seems to be revisited in this addendum. Is a per-protocol analysis still considered relevant? If not, should the Section 5.2.3 in the original ICH E9 be amended? |  |
| 560 |  | Suggest in this section, more realistic examples could be given, appropriate for each strategy, laid out in such a way which could be plausibly used in a protocol template; that would promote good practice. E.g. stating attributes of the estimand, assumptions and how they will be investigated. |  |
| 564-569 |  | Comment: Is it really a requisite to have the sensitivity analysis fully specified in the protocol? It may be reasonable to provide high level elements in the protocol and then provide technical details in the SAP? Same comment also applies to other sections of the addendum. |  |
| 564 |  | “protocol and the analysis plan”  Comment: Does this refer to two separate documents still (i.e. protocol and SAP); does this mean that the detailed descriptions would be contained within the SAP? Or does this mean that the protocol now needs to have a more robust analysis plan section (beyond describing attributes of the estimand).  Proposed change: Clarification should be provided regarding the details to be provided in the protocol versus the SAP. |  |
| 566-569 |  | Comment: Sensitivity analysis for secondary analysis may not always be required.   Proposed change: …each with a corresponding main estimator and a suitable sensitivity analysis”, suggest to change to “each with a corresponding main estimator and if appropriate a suitable sensitivity analysis. |  |
| 567-575 |  | Comment: What is meant by properly documented for estimands other than the primary? Please clarify which should be specified in the protocol and which could be left for the analysis plan. |  |
| 569-570 |  | Comment: The text suggest that even explorative analyses should be described by estimands, that seems like a lot of documentation to go into for example a protocol  Proposed change (if any): Suggest to clarify that only analyses to support claims (primary, key secondary) should be documented to the level of estimands |  |
| 579-580 |  | Comment: A statement such “Beyond these aspects, the conventional considerations for trial design, conduct and analysis remain the same. ” could appear earlier in the addendum.  Proposed change (if any): A similar statement may be introduced in Section A.1. (for example after line 70-71. It could set the stage for the whole addendum. |  |
| 580-582 |  | Comment: It is not clear when one needs to account for multiple testing in the context of additional estimands. Do you consider sensitivity analyses for missing data a source of multiplicity?  Proposed change (if any): Clarify when multiple estimands lead to a need for type I error control. |  |
| 584 |  | Comment: It is not fully clear what “pre-specified” means. Do you mean “before first patient first visit” (implemented in trial protocol or SAP). Often, “pre-specified” is interpreted as being before breaking the blind. |  |
| 585-586 |  | Comment: The added text adds more information to the sentence to help the reader better understand the intercurrent event concept.   Proposed changes: Intercurrent events that were not foreseen at the design stage but were identified during the conduct of the trial should be discussed to specify both the way the intercurrent events were handled in the analysis and the effect they had on the chosen analysis estimates and the interpretation of the trial results. |  |
| 589 |  | Comment: Section A.7. appears to include only conceptual statements indicating the interpretation of the estimand without specifying how the intercurrent event is handled in the estimators. Additional detail on this aspect would be helpful. |  |
| 593 |  | Comment: Delete this paragraph unless it can be specified how this is related to estimands. |  |
| 604-624? |  | Comment: This section reads as a non sequitur immediately following the introduction to the vignette.  Proposed change: Break this example out into its own section, analogously to Sections A.7.1 and A.7.2 for one and two intercurrent events (i.e. this one would be an example with zero intercurrent events). |  |
| 608, 636, 695, 718 |  | Comment:  Proposed change (if any): For clarity, consider changing the phrase "change from baseline to month six..." to "change from baseline at month six…" |  |
| 616-620 |  | Comment: When it comes to missing data, the addendum focuses on imputation and interpolation methods. Other methods such as MMRM are not mentioned.  Proposed change (if any): It will be helpful to include examples on what missing data methods could be applied as analyses for each intercurrent event strategy. |  |
| 617/654/789/823 |  | Comment: Is the first sentence referring to how the estimate reflects uncertainty in imputations? ”In the case of missing measurements, data need to be predicted based on plausible assumptions that account for the uncertainty due to missing data. For example, missing data may be imputed based on similar subjects who remained in the trial.”  Proposed Change replace “data need to be predicted” with “the estimate needs to be predicted” or “data need to be imputed” and add these terms to the glossary. |  |
| 617-618 |  | Comment: Minor alteration of the sentence because we think that only using “similar patients …” without further accounting in the analysis would not be a correct approach, however one which might be often incorrectly used. Therefore, we added “also”.  Proposed change: …missing data may be imputed based on also using similar subjects who remained in the trial. |  |
| 618-620 |  | Comment: “…and information on the intercurrent event”: Discrepancy to line 609 (“…no intercurrent events to be taken into account”) Please delete.  Proposed change: Similarity may be established based on the same baseline covariates, the same randomised treatment arm, and the same measurement history. and information on the intercurrent event. |  |
| 620-622 |  | Comment: This sentence suggests that “Sensitivity analyses should be pre-specified in the trial protocol”. However, in the past, sensitivity analyses were more commonly defined in the statistical analysis plan rather than the protocol.  Proposed change (if any): Change to “Sensitivity analyses should be pre-specified in the trial protocol or the statistical analysis plan”. |  |
| 620 |  | In an example of trial without intercurrent events, the guideline says: “For example, missing data may be imputed based on similar subjects who remained in the trial. Similarity, may be established based on the same baseline covariates, the same randomised treatment arm, the same measurement history and information on the intercurrent event.”.  The last reference to an intercurrent event appears a typo in this example without intercurrent events. |  |
| 626-627 |  | Comment: Move this sentence to line 611 where it makes the meaning clearer. |  |
| 629 and beyond |  | For each of the example cases it would be good to include the text description of the objective related to the estimand. This would be very helpful as a guide to include in protocols.  Proposed Change: Include text descriptions as indicated:  e.g. Estimand:Treatment policy strategy “Compare experimental drug X and placebo in terms of improving endpoint Y at  6 months for all randomized patients without regarding adherence to randomized treatment”  Estimand:Hypothetical “Compare experimental drug X and placebo in terms of improving endpoint  Y at 6 months for all randomized patients as if all patients had remained in the trial  and received treatment as planned without rescue medication until 6 months”  Estimand:Composite “Compare experimental drug X and placebo in terms of a clinical responder  at 6 months (Responder defined as achieving a pre-specified threshold of endpoint Y and not requiring rescue)” |  |
| 634-662 |  | Comment: This is a good example that makes mention of a trial design (2-arm parallel) and model (ANOVA) when all rescue therapies are considered equal and no adjustment for rescue therapy is made. This addendum/supplement would benefit greatly by also including an example where not all rescue therapies are considered equal, and the distribution of the types of rescue therapy taken differs by initial treatment group.   Proposed change (if any): Discuss adjustment for rescue therapy and SMART designs. |  |
| 640 |  | Comment: In this case, treatment policy is the wrong phrase as nobody would ever be prescribed placebo. So in this case, the true treatment effect will be attenuated due to an increased level of switching from placebo; this is the wrong message to give patients.  Proposed change (if any): Clarify the effect being measured is relating to the clinical trial but would not be reflective of clinical practice. |  |
| 648 |  | Comment: Minor alteration of the sentence for greater consistency with the content in the paragraph.   Proposed change: A similar sentence can be constructed for each of the examples strategies below,… |  |
| 652 |  | In the Treatment Policy estimand example, a possible method of analysis for obtaining the population-level summary is an Analysis of Variance. Would this be considered to be the ‘estimator’ – if so, it might be helpful to include that clarification here. |  |
| 653-656 |  | It will be helpful if Treatment Policy Strategy example is elaborated to cover the case where the data are missing after the intercurrent event |  |
| 655-656 |  | “For example, missing data may be imputed based on similar subjects who remained in the trial”  Comment: This line is redundant to Line 617-618 (“For example, missing data may be imputed based on similar subjects who remained in the trial”) Proposed change (if any):Suggest deleting one or the other. |  |
| 675 |  | Comment: Is this guidance to now collect reasons for missing data? Lack of efficacy cannot be assumed.  Proposed change (if any): Clarify the importance of collecting more specific details relating to why data may be missing and/or reasons for switching. |  |
| 682 |  | Comment: Section is highly repetitive of other sections.   Proposed change (if any): Consider aligning paragraph to reduce text |  |
| 683 |  | Can MAR assumption be used for intermittent missing values, i.e., missing data before the event? |  |
| 686-688 |  | Comment: Alteration of the sentence to be consistent with other examples  Proposed change: Similarity may be established based on the same baseline covariates, the same randomised treatment and the same measurement history and information on the intercurrent event. |  |
| 692-713 |  | Comment: For the hypothetical strategy, it would be helpful to give more details on how the underlying assumptions would be spelled out (in this hypothetical setting) and which type of sensitivity analyses could be conducted. |  |
| 698-700 |  | Comment: It should be clarified that the clinical relevance of this estimand relies on an untestable assumption: specifically, that the treatment has no impact on the effectiveness of the rescue medication. Suppose hypothetically that the treatment affects the patient in such a way that the rescue medication, which provides benefit in most situations, actually causes harm in this situation. This would be critical clinical information (since presumably the rescue medication will be used in clinical practice), but something ignored by the proposed estimand. The effect described here could lead to an average ITT effect of 0 and an average non-ITT effect > 0. If such a hypothetical situation is possible, and if the ITT effect is the more clinically relevant in that situation, then ITT should be chosen estimand. This same issue could be described for most or all of the non-ITT estimands.  Proposed change (if any): Clarify the clinical relevance of estimand using the hypothetical strategy. |  |
| 701-703 |  | Comment: In this scenario, data through month six on subjects who switch to rescue medication may still be useful to predict the measurements under the hypothetical estimand in a sensitivity analysis, depending on the assumptions. For example, one assumption may be that a subject’s response at six months under the hypothetical of not being offered rescue medication is less than or equal to the response under the scenario of being offered rescue medication. Under this assumption, the observed response at six months on rescue medication is useful in predicting the response under the hypothetical of not being offered rescue medication.   Proposed change (if any): Reword to make it clear that data collected after the intercurrent event may be useful in some cases even when a treatment policy estimand itself is not of interest. |  |
| 701-704 |  | “There would be no need to collect measurements after switching to rescue medication, unless there is interest in alternative trial objectives that would require such data (e.g. to collect safety information even after the intercurrent event)”.   Comment: Some of the structural model based approaches would require collection of these data.  Proposed change: Change to “There would be no need to collect measurements after switching to rescue medication, unless there is specific interest related to alternative trial objectives (e.g. to collect safety information even after the intercurrent event) or statistical methodology requiring the use of structural model based approaches”. |  |
| 704-705 |  | Comment: Upon a first reading, this might look like there is a typo here and it should read “is regarded as missing”. Does this imply that immaterial data that are removed (or not collected) should not be treated as missing in the resulting analysis/imputations? If data are not considered relevant then should we introduce terminology for this and add to glossary (see next comment) e.g. immaterial data (or could use alternative words such as extraneous, inapplicable, redundant, irrelevant). |  |
| 705-708 |  | We suggest to add: “A statistical analysis [..] to subjects. If values are collected after the event they must not be used in the analysis.” |  |
| 706-709 |  | Comment: Regarding the sentences "A statistical analysis for this estimand will rest on assumptions about the measurements that would have been observed under the hypothetical setting where rescue medication was not available to subjects. Generally, the assumptions needed for such predictions cannot be verified based on the observed data so that a sensitivity analysis will be necessary to assess the robustness of conclusions.", it is unclear what kind of primary analysis (i.e. estimator) is expected.   Proposed change (if any): Because this part is a general example, it would be better to include some examples of the estimator in hypothetical strategy (e.g. MMRM, MI, WGEE, etc) as in treatment policy strategy. |  |
| 707-708 |  | It will be helpful to provide some examples of prediction for the hypothetical strategy |  |
| 708 |  | Comment: We propose addition of an example to help the reader better understand the concept.   Proposed change: For example, using a Random Coefficients Model extrapolates data in a line based on observed data available for each patient. |  |
| 714-739 |  | Comment: We would appreciate clarity on how to handle subjects who have an intercurrent event, despite the population of interest being chosen to avoid this e.g., subjects are only included if no rescue medication was required during a run-in period, but a subject still takes rescue medication during the trial |  |
| 714 |  | Would it be relevant to provide an example that matches with the strategy proposed? |  |
| 718 and 746-747 |  | Comment: Lines 746-747 indicate that the intercurrent event is the use of rescue medication. However, this is not clear in the definition of estimand (see Line 718).   Propose change (if any) “B. Variable: change from baseline to month six in the designated measurement, and no switching to rescue medication occurred.” |  |
| 725-727 |  | Comment: It is not clear how a randomised withdrawal design helps in targeting “patients that would not require rescue medication”. More explanation will be helpful. |  |
| 728-39 |  | please clarify more precisely what an “appropriate” analysis is for this estimand. |  |
| 728-735 |  | Comment:  In this section, it is well explained what is not correct “A suitable analysis cannot be achieved by restricting the analysis to those subjects who did not switch to rescue medication”, but the way to obtain “an appropriate analysis to account for this confounding” is not described. An example where such analysis is possible, could help for clarity. |  |
| 730-731 |  | Comment: These individuals cannot, in general, be identified even after data collection (as mentioned in lines 257-259).  Proposed change (if any): Remove the words “in advance.” |  |
| 735 |  | Comment: The guidance should give clear steer for the reader as to what is needed. The requirement to conduct ‘an appropriate analysis’ in this section does not add sufficient detail to guide the reader in this context.   Proposed change: Please add more details about the appropriate analysis. |  |
| 740-747 |  | Comment: The terminology ‘average’ can have multiple meanings and is inaccurate in this context.   Proposed change: Correctly describe which type of ‘average’ in the three instances where the word is mentioned in this section. |  |
| 740-761 |  | Comment: Please add clarification or details on whether and how treatment duration should be accounted for. |  |
| 755-757 |  | Comment: The way to “interpolate” intermittent missing measurement is not clear. It looks like intermittent missing data may be imputed assuming they are missing at random (providing uncertainty is taken into account), but it is not clear. |  |
| 757-758 |  | Comment: “…the assumptions of the interpolation method”: add same wording as in other examples (“the assumptions of the imputation method”) to be consistent through the whole document. If there is a reason why it is not added here then please clarify.   Proposed change: the assumptions of the interpolation imputation method |  |
| 758-761 |  | Comment: It looks like a “LOCF” approach. Can it be clarified. |  |
| 762 |  | The document does not make it clear what to do if a subject has more than one intercurrent event. For example if a patient switches treatment and subsequently receives rescue medication how is this handled in the defined Estimand. |  |
| 763-764 |  | Comment: It is unclear why only discontinuations due to AE are taken into account. How would discontinuation for other reasons be considered? |  |
| 764 |  | Comment: Provide an example where all types of treatment discontinuations are accounted for or eliminate “due to an adverse event” from the second intercurrent event definition. If too complex, please acknowledge that a clinical trial would have to deal with all types of treatment discontinuations as intercurrent events (and potentially other intercurrent events). |  |
| 771-772 |  | Comment: To be consistent throughout the document, please describe all estimand strategies in this section as well. |  |
| 787 |  | Comment: Section is highly repetitive of other sections.   Proposed change (if any): Consider aligning paragraph to reduce text |  |
| 794-796 |  | Comment: It is suggested that this be clarified to indicate that these supplemental estimands could be used for estimation instead of hypothesis testing.  Proposed change (if any): “… such as contrasting the proportion and timing of rescue switchers and treatment discontinuations due to adverse events between the treatment groups. These supplemental estimands could be used for estimation instead of hypothesis testing.” |  |
| 817-20 |  | we would suggest to add that missing data methods should be used to handle missing values, or values observed after the event and hence excluded from the analysis. |  |
| 823-828 / 831-835 |  | Comment: It seems unclear why missing measurement wording is repeated in the text after the bullets.   Proposed change: Delete “In case of missing measurements, data need to be predicted based on plausible assumptions while accounting for the added uncertainty due to missing data. For example, missing data may be imputed based on similar subjects who remained in the trial. Similarity may be established based on the same baseline covariates, the same randomised treatment arm, the same measurement history and information on the intercurrent event” in one of the two places. |  |
| 826-828 |  | Comment: Please add “e.g. timing” everywhere where the following wording is used or at least when the sentence is used for the first time (line 657): “Similarity may be established based on the same baseline covariates, the same randomised treatment arm, the same measurement history and information on the intercurrent event.”  Proposed change to line (657) (first occurrence): “Similarity may be established based on the same baseline covariates, the same randomised treatment arm, the same measurement history and information on the intercurrent event, e.g. timing.”  Proposed change to line (827): “Similarity may be established based on the same baseline covariates, the same randomised treatment arm, the same measurement history and information on the intercurrent event, e.g. timing.” |  |
| 831-835 |  | Comment: “In case of missing….intercurrent events” seems to be redundant with lines 823-827 |  |
| 840/Glossary |  | Comment: For completeness, it would be helpful to add the following definitions, noting that the distinctions between variable and endpoint, and population and analysis set are clarified (as not always clear in the original E9):  “population” - the set of patients who might be exposed to the investigational treatment (which is the focus of statistical inference).  “analysis set” - the set of observed data to be included in the analysis “Endpoint” – a quantity which is derived from one or more variables (e.g. change from baseline, AUC, Cmax, time to event or censoring) observable on a single subject that directly addresses study objectives. “missing data” – value(s) that, though intended to be observed per-protocol, have not been observed “immaterial data”-data that are not considered to be relevant to the estimand which may not be measured or if measured, may be removed prior to analysis (e.g. data after initiation of rescue medication in the hypothetical estimand). “modified analysis set” - the set of observed data to be included in the analysis where immaterial data are intentionally excluded.  “subjects” is used for the participants in a clinical study to distinguish from “patients or individuals in the wider population”  “variable” - the measure which could be measured directly on any patient in the population (and which we intend to measure on the subjects in our study).  “impute” - assign (a value) to a missing data point by inference under certain assumptions (which should be specified).  “predict” –to estimate the value of an unknown quantity (e.g. the estimand). |  |
| 840 |  | Objective and variable (or endpoint) are used within R1. To have a comprehensive Glossary of terms used within this Guidance we would suggest adding Objective and Variable (or endpoint) to the Glossary. As there is often a mix between Objectives, Endpoints, and associated statistics (e.g. EMA Guidance on the evaluation of anticancer medicinal products in man regards ORR as an endpoint; standard statement in presenting results of a trial is ....endpoint met...) adding both terms might be beneficial for clarification  Proposed change (if any): Objective: Determine the scientific research questions the clinical trial should answer and will lead to defining the estimands.  Endpoint: An endpoint is an individual subject based quantitative measurement intended to reflect the effect of a drug – maybe extend by – as required by the objectives |  |
| 843-845 |  | Comment: Update text here in line with any change to the attribute wording. Proposed Change: for example, change to “Attributes of an estimand include the population, the variable(s) and endpoint of interest, the specification of how to account for intercurrent events, and the population-level characteristic defined on the endpoint of interest which is the focus of our estimation. |  |
| 847 |  | Add that the estimate is the numerical value ‘of the population-level summary for the variable (or endpoint) of interest’, otherwise this doesn’t link with the general description of an estimand, and also it doesn’t state what the ‘numerical value’ is supposed to represent in this context.  Proposed Change: Estimate: Is the numerical value of the population-level summary for the variable (or endpoint) of interest, computed by an estimator based on the observed clinical trial data. |  |
| 849 |  | Comment: It is suggested that the term “intention to treat principle” be defined as this can be misunderstood to also mean that all observations on a study subject must be used in study analysis.   Proposed change (if any): “Is the analytic approach to compute an estimate from observed clinical trial data.   Intention to treat principle:   All study subjects randomised to the study are included, in some manner, in the analysis set, and allocated to the group according to randomisation. This does not relate to how each subject’s post-baseline information is used, or what data observations are used in calculating the study result.” |  |
| 850-952 |  | Clarify whether Study Withdrawal would be considered to be an intercurrent event or not, This event would ‘preclude observation of the variable’, and therefore it would meet this definition as currently stated in the glossary. However, if Study Withdrawal is not considered to be an intercurrent event then could this definition be amended in order to clarify this? For example “Events that occur after treatment initiation and before study withdrawal, and either preclude observation…” (If that is the correct interpretation of this).  Proposed Change: Events occurring after treatment initiation and before study withdrawal, that affect interpretation of a variable. |  |
| 850-852 |  | Comment: Does the definition of intercurrent events need to be more precise, as any event resulting in missing data would fit the current definition? For example, the case of a batch of blood samples lost due to laboratory problems which result in a marker for disease not being measurable would meet this definition. However, the preclusion of these observations would affect all treatments arms equally so it should not meet the criteria to be classed as an intercurrent event. Suggest delete the phrase “preclude observation of the variable”?  Proposed Change: Change “Events that occur after treatment initiation and either preclude observation of the variable or affect its interpretation.” to  “Events, occurring after treatment initiation and before study withdrawal, that affect interpretation of a variable.” |  |
| 851-852 |  | Comment: Are events that occur after randomisation and before treatment initiation and either preclude observation of the variable or affect its interpretation considered as intercurrent events? |  |
| 867 |  | we would suggest to write “[…]same estimand but different estimators, […]”. |  |
| 872 |  | Comment: Why are supplementary analyses a “broader class” than sensitivity analyses?  Please clarify the relation between supplementary & sensitivity analyses.  Proposed Change: “The term describes a broader class of analyses than sensitivity analyses. A supplementary analysis may target a different but related estimand using the same variable(s) (e.g. repeating a composite endpoint with a different threshold) or investigates the attributes of the intercurrent event (e.g. investigating time to start rescue medication or proportion on rescue medication).” |  |
| 979 |  | It would be beneficial to add a generic example, placed in an added section (A.8), that has death as the intercurrent event. This new example can be similar to the generic example already provided (section A.7) but where the disease is life-threatening, as occurs in many oncology trials. |  |
| Processes |  | Recommendations outside the scope of a Guideline  The following recommend processes, rather that stating requirements with regard to estimands. Although good processes would help to lead to good estimands, it is questionable that the process of arriving at an estimand should be laid down in Guideline.  Lines 57-60 up to “…is the reverse” [requires that intercurrent events be considered before defining safety/efficacy variables; this may lead to efficiency but does not seem appropriate as instruction in a Guideline].  Lines 66-68 “It aims…address” addresses processes rather that what is required of an estimand.  Line 119 to “…(Figure1)” and Figure 1 itself addresses processes rather that what is required of an estimand.  Lines 130-138 and 297-308 address processes rather than attributes required of an estimand.  Lines 392-409 up to “…(Section A.5)” addresses processes rather than attributes required of an estimand. |  |
|  |  | **Text that does not support the Guideline or add substance to the Guideline**  Lines 123-4 “A suitable…selected” is an unnecessary introduction to the next sentence.  Lines 183-187 from “, e.g. the mean change…” examples of summary measures are not needed for the Guideline.  Lines 217-231 Such extended examples of estimands are suited to a tutorial paper but not to a Guideline.  Lines 335-342 describe suppositional positions of regulators, but do not provide actual guidance.  Lines 343-348 belongs in a survey/tutorial paper on estimands, rather than a Guideline.  Lines 444-449 constitute a vague warning on including multiple objectives. The issue of multiplicity is covered by other guidelines. Suggest either omit or simply refer to other Guideline(s). |  |