EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

<2018-02-28>

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

Comments from:

Name of organisation or individual
International Society for Clinical Biostatistics: Statistics in Regulatory Affairs Subcommittee

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).

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

1. General comments

Stakeholder number *(To be completed by the Agency)*	General comment (if any)	Outcome (if applicable) *(To be completed by the Agency)*
	RB: The new ICH E9 (R1) addendum on estimands and sensitivity analysis represents a mixture of useful clarifications, trivial explanations (neglecting well-known approaches of evidence-based medicine), and a number of critical issues. I recommend to revise the addendum taking the well-known PICOS approach into account and avoiding estimands which cannot be estimated without a high risk of bias and contradict statistical principles for clinical trials of the ICH E9 guideline. This requires a complete revision of the guideline, because only two of the described strategies (treatment policy, composite) should be used in general as main analysis. The other three strategies (hypothetical, principal stratum, while on treatment) are useful only as possible supplementary analysis for hypothesis generation or sensitivity analysis in special situations.	
	DSBS: In practice, linking the estimand information to the objective is tricky. For example, the generic examples do not specify what the objective is – please add that to clarify how to link the two	
	DSBS: The phrasing of an estimand under a given strategy is difficult. Suggest to add examples of how that could be done.	
	DSBS: The directional description of objective -> estimand -> design in practice is likely much more	

Stakeholder number *(To be completed by the Agency)*	General comment (if any)	Outcome (if applicable) *(To be completed by the Agency)*
	circular, suggest to reflect that in text, since the graphic illustration suggest it to be linear	
	DSBS: The addendum targets all confirmatory trials, but endpoints based on survival analysis do not really fit well in the template used. Any indication of how to use it for survival analyses would be appreciated	
	DSBS: Enrichment designs are often used to address some of key issues in clinical trials; high placebo response, high withdrawal rates. Enrichment designs are hard to fit into the template provided in the addendum	
	DSBS: An analysis that aims at testing the robustness like tipping point analyses or placebo mean imputation will not target the same estimand, as the analysis that it is testing the robustness of. This means that a lot of obvious choices for sensitivity analyses will not fulfil the requirement that the sensitivity analysis targets the same estimand. Please clarify whether sensitivity analyses can be targeting other estimands, or at least other assumptions on behaviour after withdrawal. And please provide examples of possible sensitivity analyses that targets the same estimand, to illustrate how close a match is needed on the target estimand.	
	DSBS: The concept of "conservatism" is not mentioned, suggest to mention that it remains a critical point that analyses are not set up to provide undue advantages for the new drug being tested.	
	DSBS: It would be very beneficial for the reader if the	

Stakeholder number *(To be completed by the Agency)*	General comment (if any)	Outcome (if applicable) *(To be completed by the Agency)*
	examples could be a bit more tangible. Please consider to elaborate a bit more on how the five different strategies are to be implemented e.g. regarding missing data imputation. Furthermore, it would ease the introduction of the topic 'estimands' if some of the current practices in clinical trials could be translated to estimands.	
	JS: 1. The guideline appears to reflect two different paradigms for conceptualizing an intercurrent event: a. Intercurrent events represent qualitative treatment outcomes of interest. An estimand attempts to summarize what actually happened, both classical and intercurrent events, as the complete treatment outcome including both quantitative and qualitative elements. Improved better methods better incorporate the qualitative intercurrent event information. b. Intercurrent events do not represent treatment outcomes of interest. An estimand ideally attempts to summarize a counterfactual scenario, what would have happened if the intercurrent event had not occurred. Improved methods better adjust for the qualitative intercurrent event	

Stakeholder number *(To be completed by the Agency)*	General comment (if any)	Outcome (if applicable) *(To be completed by the Agency)*
	information. While both approaches recognize intercurrent events as statistically informative rather than representing statistical noise, they lead to different directions both for study design and for methodological research. Suggest providing more guidance as to which conceptualization might apply in which circumstances. The guideline generally presents intercurrent events as potentially providing positive information. But the available methods to address intercurrent events generally appear to take a more counterfactual perspective. Suggest clearly articulating the goal separately from whether available methods reach that goal. 2. Available counterfactual approaches depend on strong assumptions, so when they are used, post-hoc checks must be made whether these assumptions remain plausible. Accordingly, the guideline very understandably focuses on sensitivity analyses. The goal for counterfactual approaches, however, is where possible more robust counterfactual methods which are less dependent on assumptions and require fewer sensitivity analyses. Usable guidance must of course help industry use currently available methods today. Nonetheless, part of the purpose of the guidance should be to help inform methodological research of what is needed. To this end, recommend clearly identifying the goal of robust methods, indicating that the goal is not generally supportable	

Stakeholder number *(To be completed by the Agency)*	General comment (if any)	Outcome (if applicable) *(To be completed by the Agency)*
	with current methods, and indicate that the purpose of sensitivity analysis is not to make them a goal in themselves, but as a necessary adjunct to methods that make strong assumptions, designed to assess the appropriateness of the assumptions made. Recommend the guideline explicitly call on the research community to help develop more robust methods where possible and appropriate. 3. Suggest additional terminology to help clarify and distinguish the goals. As one possible set of terminology in the epistemological tradition, when a counterfactual approach is used, what is observed could be called a phenomenon, while the unobservable, counterfactual estimand of interest might be called a noumenon. Different terminology could be used when intercurrent events are conceived as iintroducing additional qualitative information to, rather than being counterfactual to, the estimand of interest. 4. Trial design and methods also inform and in some cases can conceal the estimand of interest. Suggest more discussion of trial design and observation methods to reduce intercurrent events and other sources of bias. This includes explicitly evaluating compliance, patient burden, and drop-out related characteristics of methods as part of decisions about	

Stakeholder number *(To be completed by the Agency)*	General comment (if any)	Outcome (if applicable) *(To be completed by the Agency)*
	what methods to use. a. Suggest discussing general preference for simpler and less intrusive methods that may have better compliance in the context; and considering compliance and patient burden as well as ability to reliably measure an endpoint in setting visit and assessment schedules. b. Suggest discussing discontinuities in observation (e.g. clinic visits or other discrete assessment required to observe endpoint whose analysis assumes continuous observation). Issues involved can include left censoring issues (e.g. event occurs before first scheduled assessment); overestimation bias (longer observation intervals increase overestimation); etc. Dependence of visit schedule on treatment schedule can result in additional confounding (treatment resulting in longer treatment delays may appear more efficacious). Because key elements of the estimand concept, including specifying the method as part of the variable, specifying how bias will be addressed, and appropriate sensitivity analyses, are appropriate to address associated bias, integrating observation methods which introduce	

Stakeholder number *(To be completed by the Agency)*	General comment (if any)	Outcome (if applicable) *(To be completed by the Agency)*
	confounding into estimand framework and specifying and addressing observation method issues as part of the required specification process would be helpful in introducing greater rigor, reliability, and attention to sources of confounding into clinical research.	

2. 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)*
51 DSBS		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	
120 DSBS		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.	
119-123 RB		Comment: It is trivial that a clear scientific question is required before parameters are estimated. The well-known PICOS approach (participants, interventions, comparators, outcomes, and study design) should be taken into account. The given series of items goes on one hand beyond the PICOS approach (handling of intercurrent events and specification of the effect measure), but is incomplete on the other hand (intervention and comparator is missing).	

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)*
		Proposed change (if any): The given series of items should build on the well-known PICOS approach with appropriate additions.	
126-127 DSBS		Comment: Please clarify how tipping point analyses that target estimates that deviates from a model by varying measures can still be used as sensitivity analyses under the set-up in Figure 1 Proposed change (if any):	
141 DSBS		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	
151 DSBS		Comment: the figure with four bubbles used by several presenters from the addendum group, could be used to illustrate the list of the four attributes to the estimand Proposed change (if any):	
151-157 RB		Comment: In the given series of items A to D the important items	

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)*
		intervention and comparator are missing. Proposed change (if any): Add the items intervention and comparator to the described items A to D.	
154-155 DSBS		Comment: The specification C is hard to use in practice, without going into methods. For example, will MMRM "automatically" make use of all available data to influence the last observation, via the correlation, without any imputation going on. But of methods are to be kept out of the estimand specification (A-D), then C gets to be very vague Proposed change (if any): If A-D are to be void of methods, could there be some "possible methods" part where stuff like this could be described?	
182 DSBS		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 dependend 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	

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)*
210-212 RB		Comment: It is not true that the treatment policy strategy "cannot be implemented when values for the variable after the intercurrent event do not exist for all subjects". For example, imputation techniques can be used to include also subjects with missing data after the intercurrent event. Proposed change (if any): Change the statement that the treatment policy strategy cannot be implemented to the statement that the treatment policy leads to problems when values for the variable after the intercurrent event do not exist for all subjects.	
232 DSBS		Comment: The naming "Hypothetical" seems unnecessary negative. In a causal inference manner of speaking all comparisons are "hypothetical", and what makes the "hypothetical" strategy more so than the "principal strata"? Proposed change (if any): Use another term to describe the strategy, for example "Scenario" – and require that the assumed scenario is described precisely.	
232-247 RB		Comment: I question the validity and utility of the hypothetical strategy. Even if valid parameter estimation could be performed in the hypothetical scenario that an observed intercurrent event had not happened, what is the value of this estimation in practice where intercurrent events *are* happening? Moreover, no methods are available to estimate estimands in	

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)*
		hypothetical scenarios with low risk of bias. Maybe there are situations where estimands for hypothetical scenarios make sense as additional information for hypothesis generation or sensitivity analysis. Therefore, the hypothetical strategy should not be described as an option for the main analysis. Proposed change (if any): Delete the hypothetical strategy from the available options for the main data analysis. Define the hypothetical strategy as possible supplementary analysis for hypothesis generation or sensitivity analysis in special situations.	
248 DSBS		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 DSBS		Comment: A key assumption when selecting a principal stratum, is often that it should include patients that would complete the study on placebo. Such a selection may target a population with a high number of placebo responders, which could be counterproductive to showing effect in a study. Proposed change (if any): Indicate that in a number of cases principal strata may not be relevant	
248-263 RB		Comment: The "principle stratum strategy" is a purely hypothetical	

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)*
		construct. Due to the given reason (confounding) principal strata could not be formed by subsets of patients without intercurrent events. Therefore, no methods are available to deal adequately with purely hypothetical principal strata. Proposed change (if any): Delete the principle stratum strategy from the available options for the main data analysis. Define the principle stratum strategy as possible supplementary analysis for hypothesis generation or sensitivity analysis in special situations.	
264-271 RB		Comment: The restriction of the data analysis to the time period of treatment continuation leads to serious problems due to different follow-up times. Therefore, this strategy should be avoided in general. Maybe there are situations where the on treatment estimand makes sense as additional information for hypothesis generation or sensitivity analysis. However, the while on treatment strategy should not be described as an option for the main analysis. Proposed change (if any): Delete the while on treatment strategy from the available options for the main data analysis. Define the while on treatment strategy as possible supplementary analysis for hypothesis generation or sensitivity analysis in special	

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)*
		situations.	
272-276 RB		Comment: The five strategies are listed on the same level although only two strategies should be used as main analysis in practice. Proposed change (if any): Divide the list of strategies into two parts. One part with options for the main analysis (treatment policy, composite) and a subordinate part with options for supplementary analyses in special situations (hypothetical, principal stratum, while on treatment).	
272-276 DSBS		Comment: The section concerns the previous five subsections. A separate subsection for these lines could help the reader to acknowledge this. Proposed change (if any):	
302-303 RB		Comment: The formulation "Some estimands, in particular those that are estimated using the observed data, …" is unclear and makes no sense. If it is meant that an estimand is sometimes defined by observed data the statement is invalid because theoretical parameters should not be defined by observed data. If it is meant that some estimands are estimated by observed data and others not, the statement is useless, because an estimand is only meaningful if is estimable by observed data.	

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)*
		Proposed change (if any): Delete or reformulate the statement "Some estimands, in particular those that are estimated using the observed data, …".	
337-338 RB		Comment: The following statement is unclear "… but main and sensitivity estimators cannot be identified that are agreed to support a reliable estimate or robust inference." Proposed change (if any): Please clarify what is meant by the statement "… but main and sensitivity estimators cannot be identified that are agreed to support a reliable estimate or robust inference."	
464-465 RB		Comment: It is correct that "Estimation for an estimand … will require stronger and untestable assumptions if measurements are not collected following intercurrent events." Therefore, any effort should be made to collect all relevant data after occurrence of an intercurrent event. Proposed change (if any): Add the statement that any effort should be made to collect all relevant data after occurrence of an intercurrent event.	
468-471 RB		Comment: It is correct that "… the estimation of estimands constructed using a strategy that requires a hypothetical scenario to address an intercurrent event entails careful specification of	

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)*
		the hypothetical conditions and will necessarily rely on modelling assumptions that are untestable …". Therefore, the corresponding analysis should not be used as main analysis for decision making. Proposed change (if any): Add the statement that methods relying on strong untestable assumptions should not be used as main analysis for decision making.	
472-473 RB		Comment: It is correct that "… 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." If the subjects can be identified before randomisation the principal stratum strategy is nothing else than a usual subgroup analysis. If this is not the case, the principal stratum strategy can only be used as supplementary analysis but not as main analysis for decision making. Proposed change (if any): Do not use the term "principal stratum strategy" for situations of a usual subgroup analysis. In all other cases, do not describe the principal stratum strategy as an option for the main analysis.	
546-559 DSBS		Comment: The sections seems to indicate that it will not be necessary to do statistical analyses of the PP analysis set. Does this also hold for non-inferiority trials?	

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)*
		Proposed change (if any): Consider to have clear guidance concerning this	
569-570 DSBS		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	
615 RB		Comment: The method for statistical analysis is described as "… analysis of variance model with treatment group as a factor …". In the considered situation the corresponding ANOVA model reduces to the usual *t*-test. Proposed change (if any): Replace "analysis of variance model" by "*t*-test".	
682 RB		Comment: In the considered situation the use of logistic regression is not required. A simple 2x2 table with adequate statistical test would be sufficient. Proposed change (if any): Replace "logistic regression" by "2x2 table with adequate statistical test".	
692-713 RB		Comment:	

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)*
		I question the usefulness of the hypothetical setting to assume that rescue medication was not available. No regulatory decisions should be based upon such an analysis. Proposed change (if any): Describe clearly that an analysis in hypothetical settings may be used as supplementary analysis in special situations.	
724-725 RB		Comment: It is not difficult to identify members of this hypothetical population in advance; it is, in general, impossible. Proposed change (if any): Describe that it is, in general, impossible to identify members of this hypothetical population in advance and that such an analysis should only be used as supplementary analysis in special situations.	
735 RB		Comment: It is correct that "An appropriate analysis needs to account for this confounding." However, no possible methods are described, not even in an exemplary way. Indeed, no method is available which guaranties to account for all known and unknown confounders. Proposed change (if any): Add that there is no robust method available in practice to deal with all known and unknown confounders and that the corresponding analysis should only be used as supplementary	

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)*
		analysis in special situations.	
743 RB		Comment: The defined variable "average of the designated measurements while on randomised treatment" frequently leads to serious problems because the corresponding comparison is unfair due to different follow-up times. Proposed change (if any): Describe the problems of unfair comparisons due to different follow-up times and add that the corresponding analysis should only be used as supplementary analysis in special situations.	
748-750 RB		Comment: There is almost always interest in objectives requiring to collect data after switching to rescue medication. Proposed change (if any): Reformulate the statement that in general the collection of data after switching to rescue medication is required.	
802 RB		Comment: Again, the consideration of the hypothetical setting that rescue medication would not be available is useless in practice (see above). Proposed change (if any): Add the clear statement that the corresponding analysis should only be used as supplementary analysis in special	

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)*
		situations.	
813 RB		Comment: There is almost always interest in objectives requiring to collect data after switching to rescue medication. Proposed change (if any): Reformulate the statement that in general the collection of data after switching to rescue medication is required.	
842-845 RB		Comment: The important items intervention and comparator are missing. Proposed change (if any): Add the items intervention and comparator.	

Please add more rows if needed.