Assessing the Risk of Bias in Clinical Trials for Health Technology Assessments: Should Existing Tools Reflect the ICH E9(R1) Addendum on Estimands and Sensitivity Analyses?

Morga, A., Astellas Pharma Europe Ltd, London, UK
Poythress, J., Astellas Pharma Global Development, Inc., USA

HTA50

Risk of bias in the EU HTA Regulation

Article 9 of EU HTA Regulation states that the Joint Clinical Assessment (JCA) will contain a "description of both the relative effects of the health technology and the degree of certainty of the relative effects"[1]

The D4.6 EUnetHTA 21 guideline clarifies that the certainty of effectiveness results depends on: internal validity (extent to which a study is free from bias); applicability (extent to which study results provide a basis for generalisation to the target population) and statistical precision (uncertainty associated with study results due to random sampling variability)[2]

The D4.6 guideline also states that, for outcomes with evidence coming from RCTs, assess the risk of bias using the Risk of Bias 2 (RoB-2) tool[3]

The D4.6 guideline points out that the assessment of the validity of clinical studies "[..] cannot be dissociated from the population, intervention, comparator, outcome (PICO) question that will be formulated by Health Technology Assessment bodies (HTAbs)"[2].

Importantly, it is then noted that the use of RoB-2 should not exclude "the possibility of assessing evidence with an analysis strategy that corresponds best to a given PICO question, as defined according to the principles of the estimand framework outlined in ICH E9(R1) Addendum"[2].

The Risk of Bias-2 (RoB-2) tool

The RoB-2 tool[3] allows to assess the risk of bias across a number of domains, looking at the bias arising from the randomization process, the bias due to deviations from intended interventions, the bias due to missing outcome data and the bias in measurement of the outcome.

The assessment of the risk of bias with respect to deviation from intended intervention distinguishes as to whether the treatment effects of interest are represented by:

- the effect of assignment to the interventions at baseline, irrespective of whether the interventions are received as intended (the 'intention-to-treat effect') [3]; or
- the effect of adhering to the interventions as stated in the trial protocol (the 'per-protocol effect') [3]

The ICH E9 (R1) Addendum on Estimands and Sensitivity Analysis in clinical trials[4], adopted in 2019 and implemented by many Health Authorities worldwide, introduces the estimands framework and, with it, a transparent manner for the description of the treatment effects in a given trial.

The clinical research question of interest is translated in a specific range of strategies through which post-randomisation (or intercurrent) events (ICEs) are addressed.
It is therefore important to address the issue of whether the possible strategies for handling ICEs as described in the Addendum can be reconciled with the treatment effects described in the RoB-2 tool.

Ascertaining this point is important in light of the expected use of the tool in the context of the EU HTA Regulation.

ICH E9 (R1) Addendum

Per ICH E9 (R1) Addendum, an **estimand** is defined by 5 attributes:

- **Treatment**: the treatment condition of interest and, as appropriate, the alternative treatment condition(s) to which comparison will be made
- **Population**: the population of patients targeted by the clinical question of interest (e.g., as defined by the protocol's I/E criteria)
- **Variable (or endpoint)**: the measurement/assessment to be obtained for each participant that is used to address the clinical question of interest.
- **Intercurrent Events (ICEs) and strategies for handling them**: events occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest
  - each expected ICE should be accompanied by a corresponding strategy for handling the ICE that is consistent with the clinical question of interest
  - the strategy for handling an ICE precisely defines the treatment effect to be estimated
- **Population-Level Summary**: provides a basis for comparison between treatment conditions (e.g., difference in means, odds ratio, etc.)

ICH E9 (R1) discusses 5 **strategies for handling ICEs** to align with the clinical question of interest:

Strategy	Description
Treatment Policy	the occurrence of the ICE is considered irrelevant in defining the treatment effect of interest; the values of the variable are used regardless of whether the ICE occurs
Hypothetical	a scenario is envisaged in which the ICE would not occur; the value of the variable to reflect the clinical question of interest is the value that the variable would have taken in the hypothetical scenario defined
Composite	the ICE is considered to be informative about the participant's outcome and is therefore incorporated into the definition of the variable
While-On-Treatment	response to treatment before the occurrence of the ICE is of interest
Principal Stratum	the clinical question of interest relates to the treatment effect only within the population in which the ICE would/would not occur

Example of Misalignment Between the RoB-2 tool and the ICH E9 (R1) Addendum

Difficulties in assessing risk of bias arising due to misalignment between the RoB-2 tool and ICH E9 (R1) can be illustrated with a generic example of a clinical question of interest addressed via a composite strategy.

Generic Clinical Question of Interest: Does the intervention improve <*some continuous measure of response*> by <*some clinically meaningful threshold*> at <*some predefined post-baseline evaluation period*>, without the need for prohibited medications or alternative therapy?

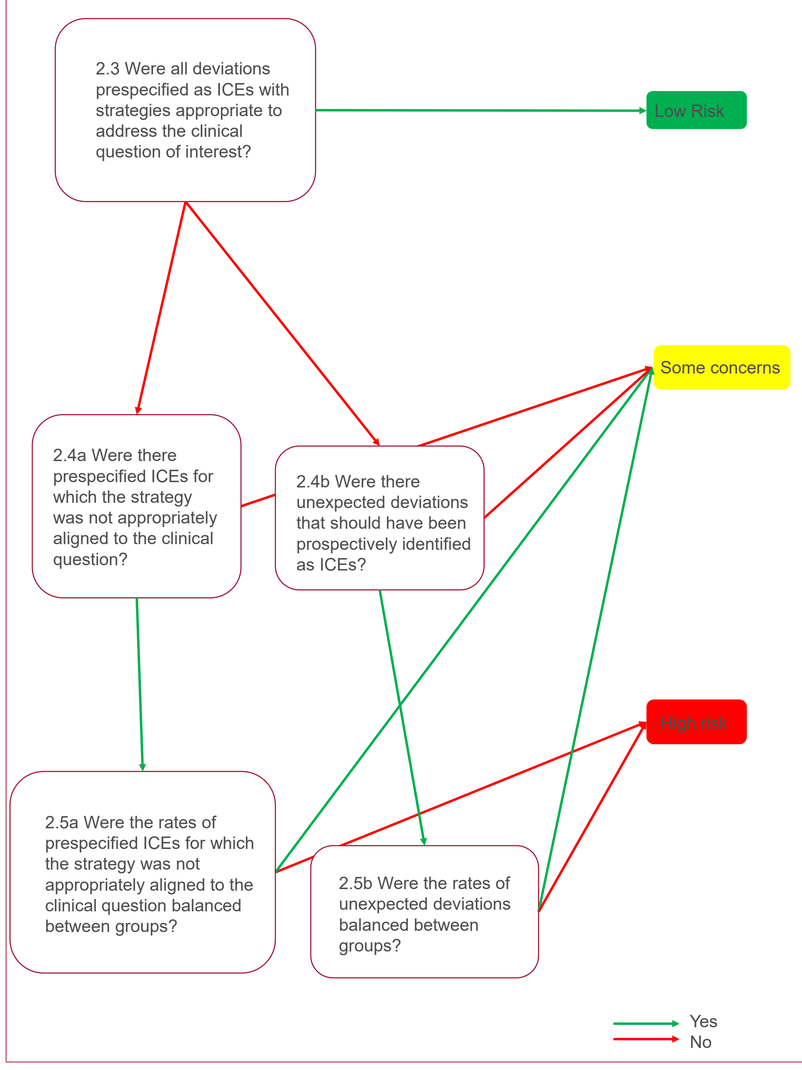
Some attributes of an estimand that could address the clinical question of interest:

- **Possible ICE**: Participant switches to alternative therapy or otherwise prohibited medication prior to the evaluation period
- **Composite Strategy** (via the endpoint/variable attribute): Nonresponse is defined as the <*continuous response variable*> does not attain the <*clinically meaningful threshold*> at the evaluation period OR participant switches to alternative therapy prior to the evaluation period OR participant takes prohibited medication prior to the evaluation period
- **Population-level summary**: Difference in proportion of responders between the interventional arm and comparator arm (i.e., difference in response rates)

Assume the ICE occurs in some participants. If the user selects to assess the effect of assignment to intervention in the RoB-2 tool, the signaling questions in the following table may be answered:

Signaling questions from RoB-2 tool[5]	Interpretation of the elaborative text	Potential in the current tool for ambiguity of question and/or variability of user response	Potential improvement if RoB-2 tool was aligned with ICH E9 (R1)
2.3 Were there deviations from the intended intervention that arose because of the trial context?	The user should answer "yes" or "probably yes" when there is evidence that the protocol-specified interventions were not implemented because of the trial context. The user should also answer 'yes' or 'probably yes' if the trial context led to the **use of interventions prohibited by the protocol**. On the other hand, the user should answer "no" or "probably no" when the changes from the protocol-specified intervention are deemed consistent with clinical practice. Similarly, the answer should be "no" or "probably no" when the changes from the protocol-specified intervention are in line with other provisions specified in the protocol, such as use of allowed concomitant medications or treatment discontinuation due to safety events.	The user should select 'Yes' based on the literal interpretation of the instructions for how to answer if trial participants used medications or alternative therapies prohibited by the protocol. However, some users may select 'No' because an appropriate strategy was chosen to handle the ICE and was prespecified in the protocol.	The critical question should be: **Were all deviations prespecified as ICEs with strategies appropriate to address the clinical question of interest?** For a prespecified ICE, it is less important whether or not it occurred, and more important that an appropriate strategy was chosen to handle it. If all stakeholders agree that the answer is 'Yes', then the results should be assessed as low risk of bias.
2.4 Were these deviations likely to have affected the outcome?	If deviations from the intended intervention are not expected to affect the outcome, then neither are they expected to affect (i.e., bias) the estimate of the intervention's effect.	Some users may select 'Yes' because the ICE precludes observation of the continuous component of the endpoint, and thus affected that component of the outcome. Other users may select 'No' because the ICE was incorporated into the endpoint attribute using a composite strategy, so the ICE itself became part of the outcome.	A common terminology consistent with the strategies discussed in ICH E9 (R1) can help avoid confusion, misinterpretation, and ambiguity. It can also help focus attention on the critical question of whether the strategy was appropriate to address the clinical question of interest.
2.5 Were these deviations from intended intervention balanced between groups?	Even if deviations from the intended intervention are expected to affect the outcome, if the rates of deviations are balanced between groups, then they may not bias the estimate of the intervention's effect.	If ICEs are handled by a composite strategy, results for the individual components of the composite variable may not be available. If results are available for individual ICEs, some users may interpret imbalance in rates of ICEs between groups as contributing to differential response rates, while other users may interpret it with respect to the effect on the continuous component of the responder endpoint.	If all ICEs were prespecified and handled with appropriate strategies, then balance between groups may not be relevant. Only imbalance in unanticipated deviations or ICEs not handled with an appropriate strategy may lead to a high risk of bias. For a composite strategy, imbalance in the rates of ICEs can be interpreted as contributing to differential response rates.

Potential revision to RoB-2 signaling questions to align to a common framework and terminology as ICH E9 (R1):



Conclusions

- The RoB-2 tool is not aligned to the estimand framework defined in the ICH E9 (R1) Addendum, particularly for identification of ICEs and strategies to handle them that are appropriate to address the clinical question of interest.
  - The 'intention-to-treat' and 'per-protocol' effects assessed by the RoB-2 tool do not represent precisely the same concepts as treatment policy and principal stratum strategies
  - It is not clear how to use the RoB-2 tool to assess risk of bias in results pertaining to clinical questions addressed by hypothetical or while-on-treatment strategies
  - As illustrated by the example for composite strategy, the misalignment in terminology between the RoB-2 tool and ICH E9 (R1) can result in ambiguity and/or variability in interpretation across users with respect to answering the tool's signaling questions.
- Use of the RoB-2 tool to assess risk of bias for randomized controlled trials as recommended in the EUnetHTA 21 methodological guidelines has the potential to lead to incorrect assessment of risk of bias and/or unpredictable assessment outcomes in the context of the JCA
- As Regulatory Authorities around the world require clinical trial protocols and statistical analysis plans to include estimands, HTA bodies should rely on a tool that is aligned with the ICH E9 (R1) Addendum, ensuring that conclusions from the use of the tool are transparent and robust. Methods used by HTA bodies to evaluate trials for risk of bias should be standardized, predictable and unambiguous to all stakeholders.
- It is therefore recommended that the Cochrane RoB-2 tool is updated so to reflect the ICH E9 (R1) Addendum.

References

1. Regulation (EU) 2021/2282 of the European Parliament and of the Council of 15 December 2021 on health technology assessment and amending Directive 2011/24/EU, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021R2282 (accessed October 2023)

2. EUnetHTA 21 - Individual Practical Guideline Document, D4.6 Validity of clinical studies, Version 1.0, 16.12.2022, available at: https://www.eunethta.eu/wp-content/uploads/2022/12/EUnetHTA-21-D4.6-Practical-Guideline-on-validity-of-clinical-studies-v1.0-1.pdf; accessed October 2023

3. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane, 2023; available from www.training.cochrane.org/handbook; accessed October 2023

4. International Council of Harmonization. Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials E9(R1); 2019. available at: https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf; accessed October 2023

5. Revised Cochrane risk-of-bias tool for randomized trials (RoB-2): supplementary material, available at: https://www.bmj.com/content/bmj/suppl/2019/08/28/bmj.l4898.DC1/stej049848.ww1.pdf, accessed October 2023