Rethinking Censoring and Censoring Mechanisms in Light of the Estimands Framework

Jonathan Siegel, Bayer

Stefan Englert, Abbvie

Hans-Jochen Weber, Novartis

The Pharmaceutical Industry Working Group on Estimands in Oncology

A Scientific Working Group of the ASA Biopharmaceutical Section

Science For A Better Life

BIOP Virtual Statistics Workshop

September 24, 2021

Disclaimer

The views presented in this presentation are solely those of the authors on behalf of the Pharmaceutical Industry Working Group on Estimands in Oncology, Censoring Mechanisms Subteam, and do not represent the views of their respective employers or any other organization.

Outline

- Background
- Noumenal and phenomenal censoring
  - Implicit noumenal censoring
  - How should implicit noumenal censoring be interpreted?
- Design from the beginning
- The discreteness of assessments
- Occlusion
- Addressing conflicting needs
- An example
- The 5 strategies considered
- Study design: Feedback loop and engineering process
- Key recommendations
- Summary and conclusions
- References

Background: Estimand Guidance

- This presentation assumes familiarity with the ICH E9 (R1) Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials.
- Oncology trials rely heavily on time-to-event variables, especially in late-phase regulatory approval trials
- The estimands framework is of particular importance in a time-to-event context.
- The Pharmaceutical Industry Working Group on Estimands in Oncology Censoring Mechanisms Subteam has been meeting since 2018 to clarify applicability of estimands guidance in time-to-event cancer trials.
- The team is drafting two white papers for publication, introducing concepts and applications summarized in this presentation.

Background: One size does not fit all

- In past, methods in pharmaceutical oncology time-to-event clinical trials have been highly standardized
  - A small set of methods, such as Kaplan-Meir, Cox, and log-rank tests, have been used across the board.
  - Underlying assumptions, such as non-informativity and proportional hazards, have often gone unchallenged
  - Standardized censoring rules have been applied without regard to context or the underlying clinical question
- The estimands guidance focuses studies on addressing a clinically meaningful, feasible research question
  - It requires justifying assumptions and opens a wider range of tools.
  - It has analogies to the concept of quality as fitness for purpose (Deming, 1986)

Background: The treatment policy strategy

- The treatment policy strategy comes closest to the "ITT" approach associated with Fleming (2009) and others.
  - It has accordingly gotten the most attention and deserves special mention.
- In the original context of treatment, the scientific question looks at the effect of randomization including the randomized treatment and all subsequent therapies (Carroll, 2007)
- More generally, the scientific question is concerned with the outcome from randomization to the event of interest, through and beyond the intercurrent event.
  - The strategy "sees through" the intercurrent event.
- Implementing this strategy requires an ability to consistently follow patients through and beyond the intercurrent event.
  - This is not always feasible.

New Concepts

- The Censoring Mechanisms Subteam has developed new concepts which we believe help explain estimands in the complex environment of oncology TTE studies.
  - Noumenal vs. Phenomenal
  - Occlusion

Noumenal and phenomenal censoring

- Noumenal censoring is intentional, systematic censoring done as part of the study design
  - Noumenal censoring is a property of the estimand
  - Noumenally censored data lie outside the frame of inference (actual or hypothetical) to which the analysis is inferring.
- Phenomenal censoring is censoring that occurs in an individual patient incidental to trial conduct
  - Things like individual decisions to withdraw from the trial
  - Phenomenal censoring is a property of the estimator
  - It does not change the frame of inference

Implicit Noumenal Censoring

- Protocols often end assessments at predefined events.
  - Complete safety assessments often end a fixed number of days after last treatment
  - Clinic visits and related efficacy assessments often end at radiological progression.
  - Everything ends at the end of the study.
- When assessments are systematically ended or data is discarded beyond an event, Implicit noumenal censoring occurs.
  - Based on anything in the study design (visit schedule, withdrawal criteria, even data handling rules)
  - This is very common in oncology studies
- We believe the presence of noumenal censoring, even when implicit, changes the strategy and the interpretation of the results.

How should implicit noumenal censoring be interpreted?

- When implicit noumenal censoring occurs, we generally recommend interpreting as a hypothetical strategy.
  - A hypothetical strategy asks what would have happened if the occluding event had not occurred.
  - Using censoring to implement this strategy assumes the intercurrent event is non-informative, i.e. does not change future risks.
  - Most events that end assessments in cancer trials have at least some potential for being informative
- An independent-causes while on treatment strategy (Unkel, 2019) is also possible.
  - We believe this makes sense only when there is strong reason to believe a systematically censoring event is non-informative

Independent-causes while-on-treatment strategy example: The end of study

- Consider a trial comparing an immuno-oncology product with a delayed effect to a standard constant-effect chemotherapy.
  - Hazards are non-proportional.
  - Standard estimates of the treatment effect (e.g. HR) will depend on when the trial ends.
  - Ending the trial early tends to underestimate the treatment effect compared to following all patients until event.
  - But censoring from end of trial is uninformative (past hazards are different, but not future ones)
- In this trial, censoring for end of study means standard estimators like HR are biased when interpreted as a treatment policy strategy (following all patients until the event of interest.)
- But interpreting as a while-on-treatment ("while during study") strategy means no bias (we answer exactly the question posed)

The recommended approach: Design it from the beginning

- Never leave noumenal censoring implicit.
- For each estimand, the research question of interest should be formulated, and the likely intercurrent events and strategies for addressing them considered.
- Visit schedules and withdrawal criteria should be designed with the needs of the estimands in mind, in tandem with the planned estimands and intercurrent events strategies.
- Every systematically censoring event should be explicitly identified in the estimand definition and assigned a strategy.
- Alternative strategies to censoring and calling it a hypothetical strategy should be considered.

A special problem in oncology: The discreteness of assessments

- Many oncology assessments, can only be assessed at substantial intervals apart
  - Tumor assessments (too much radiation exposure is unethical)
  - Others can only be made at clinic visits
- The occurrence of an intercurrent event can be separated in time from a withdrawal decision
- Example:
  - Patient experiences an AE between clinic visits
  - Withdraws based on that AE at the next clinic visit
  - Continues safety follow-up for some further period (e.g. 30 days)
  - This results in a time lag between the intercurrent event inducing withdrawal, and the withdrawal itself, with observations in between.
- We introduce a 2nd concept that helps address this discreteness issue

Occlusion

- Occlusion occurs when data beyond an event is not included in the analysis. (not collected, or discarded)
- It can be handled not only through censoring, but also through e.g. a competing risk event, composite event, or causal inference method.
  - It is a generalization of censoring
  - It is also a broader term than terminal event. Data can be excluded from an analysis even if collected in the trial.
- Under this approach, censoring is a particular method of implementing occlusion, one of multiple possible.
  - Occlusion permits discussing alternative strategies
- It is not the same thing as an intercurrent event
  - In the AE example above, the intercurrent event (the AE) can occur before the occlusion (the withdrawal)
  - In the end-of-study example, the end of the study is occluding, yet not intercurrent.

Occluding and non-occluding strategies

- Three of the five strategies discussed in the estimands guidance handle an intercurrent event through occlusion:
  - A while-on-treatment strategy asks what happened up until the time of the event
  - A hypothetical strategy generally asks what would have happened if the event had not occurred
  - A composite strategy incorporates the intercurrent event into the event of interest.
- All three strategies do not look beyond the event.
- Systematic occlusion by design is noumenal. The 3 occluding strategies involve noumenal occlusion, excluding what happens beyond the event from the noumenon.
- Two strategies, treatment policy and principal component continue to look beyond the intercurrent event
  - They involve only phenomenal occlusion.

Example: Handling withdrawal in TTE estimands

- In the example given above, (intercurrent event, then assessment, then actual withdrawal from assessments) we recommend:
  - For treatment policy and principal component strategies, the intercurrent event should be ignored and all assessments included
  - For a composite strategy, the intercurrent event becomes an event of interested, dated as such
  - For a while-on-treatment or hypothetical strategy:
    - If the occurrence of the intercurrent event is potentially based on subjective judgment, the date of actual withdrawal should be used and all assessments included
    - If the intercurrent event is not a matter of subjective judgment, occlusion can be dated at the intercurrent event and subsequent assessments discarded
- Reasoning is discussed in our 2nd paper.

Addressing conflicting needs

- Oncology clinical trials often have to address conflicts among different requirements
  - Dosing, patient management, and ethical-care needs often drive visit schedules and assessment termination requirements
  - Higher-priority estimands often need to drive what little flexibility is left.
  - Secondary estimands often have to accept a visit schedule with assessments terminated based on the needs of patient management and the primary estimand.
- Example: In a study with primary endpoint PFS, all clinic visits usually end at radiological progression.
  - In this case, radiological progression occludes all secondary endpoints requiring clinic visits.
- For example, time for forced lung capacity deterioration
  - We discuss alternative strategies for this example.

The while-on-treatment strategy

- The research question is only concerned with what happened up to the time of the occluding event.
  - This could be a good fit for our example. The questions may not be the ideal one we want to know, but it's a question that this study, with its constraints, can potentially answer.
- The term "while-on-treatment" might not be ideal here. It could be called "while progression-free" in this context.
  - In general, "while prior to occlusion" or "while un-occluded"
- Radiological progression is likely to be informative of deterioration.
- The subdistribution hazards approach (Fine and Gray, 1999) does not assume causal independence.
- However, dependent hazards issues limit valid use. The approach may be most appropriate when:
  - A descriptive analysis addressable by the CIF is appropriate
  - The occluding event is objectively determinable.

The composite strategy

- When an occluding intercurrent event is highly correlated with and/or clinically relevant to the event of interest, a composite strategy might be considered.
  - We posit a correlation in the example.
- It might be reasonable to consider a composition of events like time to symptom deterioration with assessment-ending intercurrent events like progression (or death).
  - It has a clear, valid interpretation.
  - It permits modeling and testing without the Fine-Gray approach's validity problems.
- When an estimand's lower priority results in assessments being occluded based on other considerations, it could be considered.
- As before, the research question it addresses is not necessarily ideal, but is valid and interpretable within the constraints of the trial.

Hypothetical Strategies

- Simply censoring for end of assessments induces an implicit hypothetical strategy when assessments systematically end following an intercurrent event
  - Making this strategy explicit, while not ideal, would represent an improvement over leaving it implicit.
- In some cases, causal hypothetical strategies could be considered addressing what would have happened if the event did not occur under a more sophisticated model than simple censoring
  - But one with additional assumptions.
- Hypothetical strategies are discussed in more detail in the Treatment Switching Subteam's presentation (and in our 2 white papers)
- In this particular example, we may not be interested in knowing what patients' QoL would have been if they hadn't progressed. But in other contexts the question posed may be more appropriate.

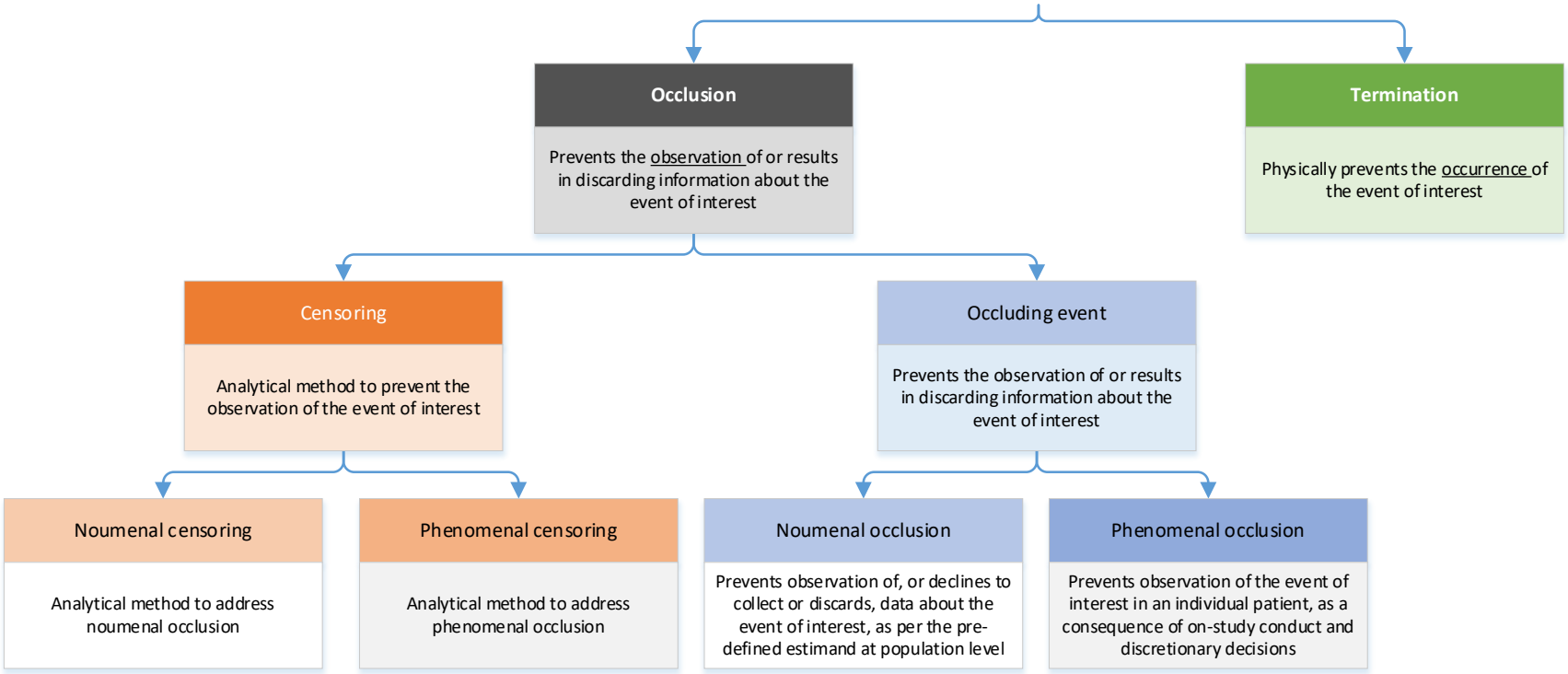
The two non-occluding strategies

- Both the treatment policy strategy and the principal component strategy do not reflect (or induce) noumenal occlusion
- All censoring in both strategies is phenomenal, based on individual events that occur in the course of the study resulting in loss to follow-up.
- Treatment policy strategy was discussed above
- The principal component strategy attempts to include only patients in whom the relevant intercurrent event is expected not to occur, based on a model of baseline characteristics.
  - Occurrences of the occluding intercurrent event notwithstanding the model are assumed not to be systematic
  - Strategy has not often been used in oncology pharmaceutical research.

Summary of Strategies for Intercurrent Events

Irrespective of	Include in Outcome	Scenario in which event does not occur	Prior to occurrence	As part of target population definition
• Outcome after intercurrent event is still of interest • Data should be collected after intercurrent event	• Define composite endpoint including the intercurrent event • Intercurrent event is informative for effect of interest	• A scenario is envisaged in which the intercurrent event would not occur	• Scientific question is about what happened prior to the intercurrent event • Outcome after intercurrent event is considered irrelevant	• Population is defined by those in whom the intercurrent event would or would not occur
Treatment Policy	**Composite**	**Hypothetical**	**While on Treatment**	**Principal Stratum**

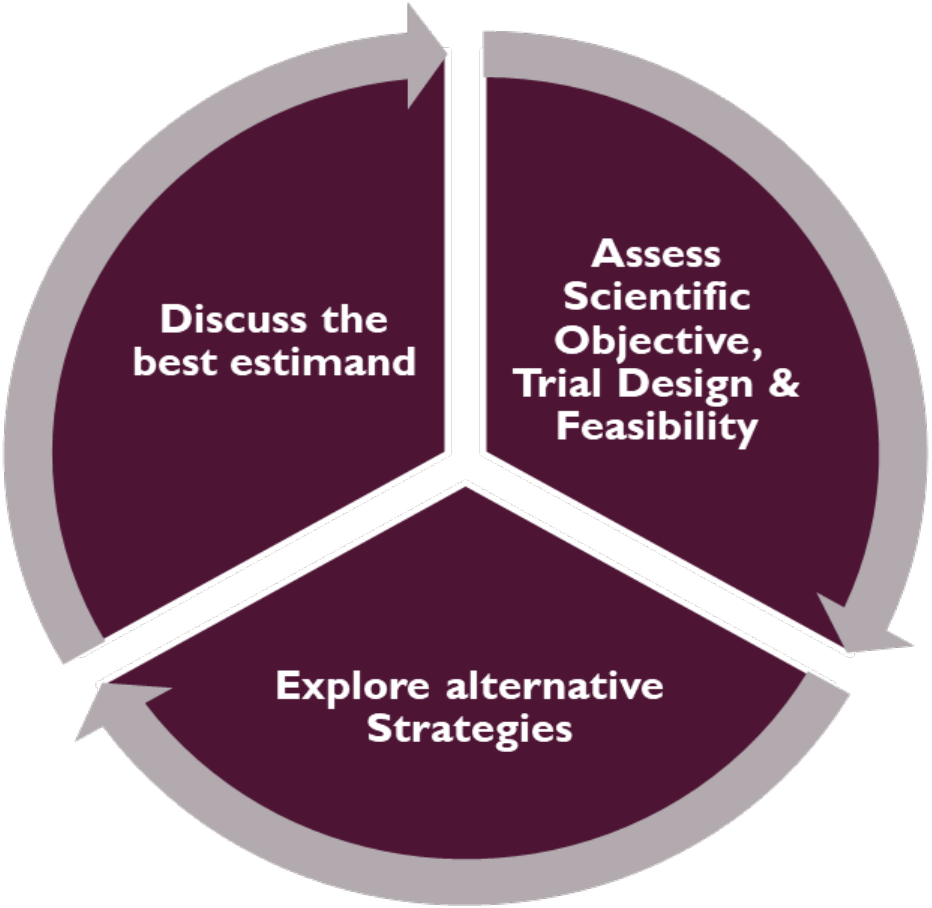
Diagram of concepts



Study design as a feedback loop

- In practice, feasibility and ethical considerations, the needs of other estimands, and other factors will often require revisiting what can be feasibly addressed in the study context.
- As a result, the design process will often involve a feedback loop analogous to Deming's Plan-Do-Study-Act cycle (Deming, 1986).
- It is important to start with the ideally desired research question, not just what appears feasibly implementable.
  - It is always possible that design changes, technology, and other improvements may overcome constraints and enable the original research question to be addressed.
  - Perhaps a member of the team will have an idea
  - Perhaps in a future trial.

Study design as a feedback loop



Towards an engineering process

- Under an estimands framework, study design requires creative collaboration among clinicians, statisticians, and study management and operations personnel.
- Clear understanding of goals.
- Clear understanding of constraints and obstacles
- Clear understanding of how the constraints and obstacles, and the strategies to address them, affect the meaning, interpretation, and usefulness of the results.
  - Fitness for purpose
- Creative solutions to overcome constraints and/or identify a compromise solution as close as feasible to the original goal.

Key recommendations

- Ensure that strategy is aligned with protocol and visit schedule
- Ensure that implementing the strategy, including the needed data collection, will be feasible in the study context
- Identify all potentially intercurrent events that systematically end assessments or remove subsequent data from analysis
- Develop an appropriate strategy to address noumenal occlusion
- Careful planning, CRF design, and training is needed to minimize both missing data and intercurrent events.
- Collect data on intercurrent and occluding events to identify and assess informativity (e.g. whether reason for treatment withdrawal was treatment-related).
- Perform sensitivity analyses to check if key assumptions underlying the strategy were met
- Consider an alternative strategy if key assumptions appear infeasible up-front or were not met post-hoc.

Summary and conclusions

- The estimands framework requires rethinking long-standing habits in time-to-event oncology trials
  - Careful consideration of research goals and strategies
  - Understanding context, constraints, and alternatives.
- Censoring is only one possible approach to occlusion, and may not be the right one.
- Noumenal occlusion (systematicly not looking beyond an event) is inconsistent with a treatment policy strategy and censoring may be an inappropriate approach to address it.
- The complex clinical context of oncology, with non-proportional hazards, correlations among outcomes, discrete assessments, ethical and patient management requirements, etc., makes designing and interpreting time-to-event estimands particularly challenging.
- Cross-discipline cooperation and an engineering approach is recommendedd.

References

- Carroll KJ (2007): Analysis of progression-free survival in oncology trials: Some common statistical issues. Stats Med 6: 99-113
- Deming, WE. Out of the Crisis. MIT Press (1986)
- Fine J, Gray R (1999); A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 94:496-509
- Fleming, T, Rothmann, M and Lu H. (2009) Issues in using progression-free survival when evaluating oncology products. J Clin Oncol 27:2874-2880.
- Unkel, S. et al. (2019). On estimands and the analysis of adverse events in the presence of varying follow-up times within the benefit assessment of therapies. Pharmaceutical Statistics 18:166–183.

Thank you!