Estimands in Oncology: A Topical Panel Discussion

Steven Sun, Jay Zhao, Jonathan Siegel, Yi-Ting Chang

Co-Chairs: Yeh-Fong Chen and Qing Xu

Panel Discussion, BIOP 2022 Conference, September 21, 2022

Disclaimer - General

- The comments expressed herein are the authors' own and should not be interpreted in any way as representing their respective employers' views or policies.

Disclaimer - FDA

The comments expressed herein are the authors' own and should not be interpreted in any way as representing FDA's views or policies.

Questions for the Panel

- 1) How can we encourage consistent analysis and interpretation of Duration of Response and Time to Response in clinical trials?
- 2) What is the clinical question of interest if patients receive the option of subsequent therapy?
- 3) How does concern about causal estimands impact the way we do time to-event trials?
- 4) What do we mean by follow-up time in a clinical trial?

Question 1

How can we encourage consistent analysis and interpretation of Duration of Response and Time to Response in clinical trials?

DoR-TTR Subteam to review DOR and TTR from the perspective of estimand framework

....

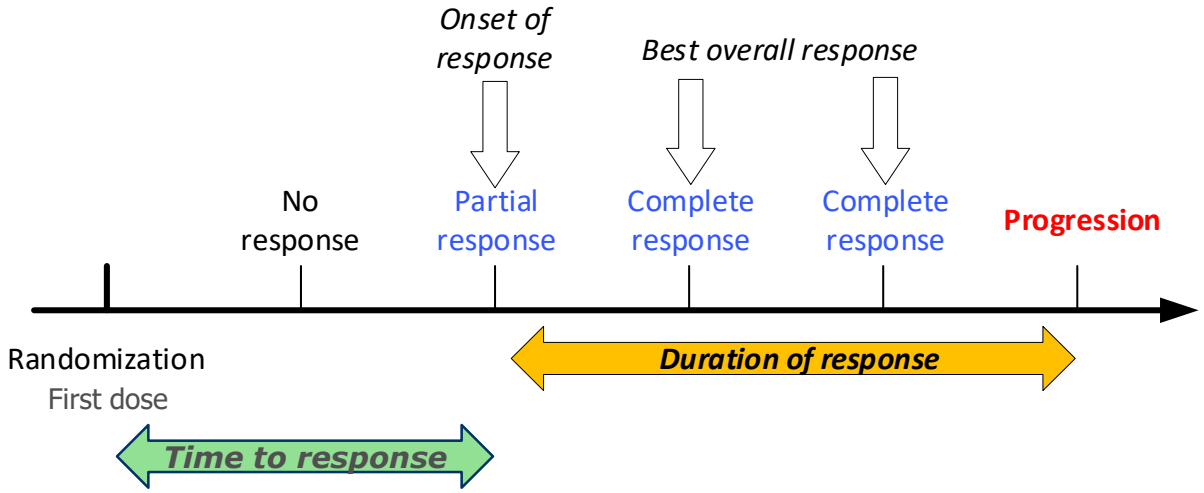
- Hans-Jochen Weber (Novartis)
- Alexander Todd (AstraZeneca)
- Jiang Li (Beigene)
- Francois Mercier (Roche)
- Oliver Sailer (Boehringer Ingelheim)
- Satrajit Roychoudhury (Pfizer)
- Stephen Corson (Phastar)
- Godwin Yung (Genentech)
- Steven Sun (Johnson & Johnson)

relevant **clinical questions**

Analytic **approaches** to quantify DOR and TTR

When and how should we **present** DOR and TTR

Response related endpoints



Duration of Response (DOR)

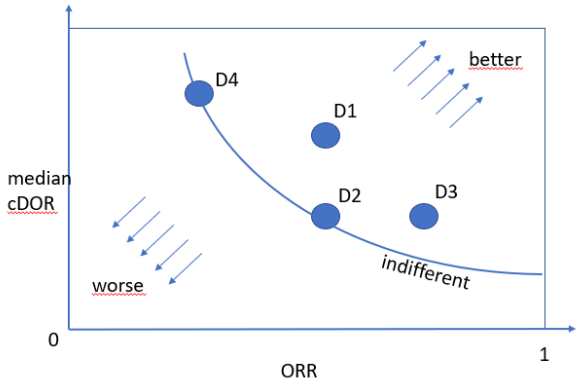
- Characterization of quality of response
- ***Always*** reported in label for AA
- Estimand often not described in SAP or publication
- **Conditional** on being a responder

Time to Response (TTR)

- Clinical interpretation: more favorable if TTR is shorter??
- Reported in label ***sometimes***
- Estimand often not described in SAP or publication
- **Conditional** on being a responder

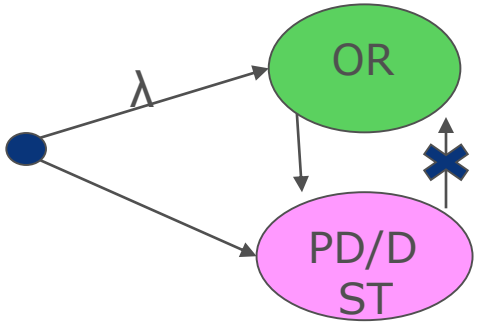
Analytic Approach

DOR



- Combine ORR and cDOR in a single **unconditional mean DOR** (**EDOR**)
- Analyse via **probability of being in response function (PBRF)** (Ellis 2008 CCT 29 456-465)
  - Area under PBRF = EDOR
- Patient level: unconditional DOR=0 if non-responder, else =cDOR
- EDOR more informative than ORR + cDOR? (Huang et al. 2020 Ann Intern Med 173: 368-374)

TTR



- Censoring at ICE of PD, death or subsequent therapy leading to **biased** estimate because these are competing risk
- Censoring at infinity/after last event date assumes that PD/death or ST time and OR time are independent – is this realistic?
- Principal stratum analysis: TTR for patients who don't have PD/D/ST – Does such subpopulation exist?

Question 2: What is the clinical question of interest if patients receive the option of subsequent therapy?

Jay Zhao, FDA
Qing Xu, FDA

2022 ASA Regulatory-industry Statistics Workshop, Rockville, MD
September 21, 2022

Case Study-Different Clinical Question Answered

Sample size - N=320
Randomization - 1:1 ratio to experimental arm and placebo arm
HSCT - 48% in experimental arm vs 20% in placebo arm

OS endpoint , Imbalanced HSCT distribution	P-value	HR (95% CI)
Primary Analysis: Cox PH model	0.02	0.75 (0.58, 0.96)
Supplement Analysis: MSM model	0.65	1.06 (0.82, 1.35)

Qing Xu, Donna Przepiorka, Pharmaceutical Statistics, 2021 Nov;20(6):1088-1101. doi: 10.1002/pst.212

Science For A Better Life

Question 3: How does concern about causal estimands impact the way we do time to-event trials?

Jonathan Siegel

BIOP 2022 – Rockville, MD – September 21, 2022

Causal Estimands and Time-to-Event Trials

- Conventional methods for estimation and testing in time-to-event trials are often not causal estimands under
  - non-proportional hazards (NPH)
  - competing risks
  - intercurrent events (ICE)
- Dependence on censoring patterns can make estimators dependent on trial design and patient behavior
- Non-proportional hazards are increasingly common in oncology

Estimands and Feasibility

- Must be clinically meaningful and feasible.
- Treatment policy strategy often preferred, not always feasible.
  - Requires consistently following patients to event of interest
- Scheduled clinic visits after progression may be infeasible
- Examples:
  - Subsequent therapy/new trial has different visit schedule
    Primary endpoint is not progression or OS but patients stop clinic visits at progression
  - Ending clinic visits at progression informatively censors all other secondary variables that depend on clinic visits
    - E.g. clinically assessed symptom indicators
- Alternative strategy should be considered
  - Treatment policy might still be the best strategy

Imperfect Alternatives: While on Treatment

- Event of interest assumed impossible/irrelevant after ICE
- RMST - causal estimand under NPH
  - Truncation perhaps analog of while on treatment
  - Operating characteristics assume specific hazard pattern
  - Absolute difference problematic. Hard to interpret
    - 3 vs.15 months ≠ 8 vs. 9 years
  - Results depend on trial design, censoring pattern.
- Fine-Gray – competiing risks - Immortal time following ICE
  - Lacks 1-1 relationship with cause-specific hazards.
  - Subdistribution HR not clinically interpretable.
- Cumulative Incidence Function preserves causal estimand
  - Descriptive only. Tests for comparing CIFs proposed but not widely received (e.g. Aly, Kaucher, and Mckeague, 1994; Zhang and Fine, 2008)

Imperfect Alternatives Cont.

- Composite strategy combines intercurrent event with event of interest. Example: PFS
  - Often least problematic strategy for addressing ICEs
  - Not necessarily clinically meaningful
- Principal Stratum identifies a population not susceptible to ICE.
  - The population for which causal estimation is valid is not necessarily the population of scientific or medical interest.
  - Modeling the principal stratum may not be reliable.
- Hypothetical strategy addresses what would have happened if intercurrent event had not occurred
  - Modeling generally dependent on strong assumptions
  - Discussed in subsequent therapy discussion.

Thoughts for Discussion

- The estimands framework requires paying close attention to assumptions, trial design, intercurrent events, and hazard patterns
- Few good solutions. New developments often yield:
  - Estimand that is causal/valid but not clinically interpretable
  - Reducing one assumption dependence by introducing others.
- Progress requires better collaboration between methodologists and clinical trialists
  - More attention to real-world problems and clinical meaning when constructing methods
  - More attention to statistical issues by clinical trial community
  - Communication and shared understanding
- Importance of balance between scientific validity, reliability, and operational feasibility

Question 4: What do we mean by follow-up time in a clinical trial?

Yi-Ting Chang, AstraZeneca

ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop

September 21, 2022

Problem statement

“Follow-up quantification”:

- ❑ Unclearly defined concept.
- ❑ Different quantities used to “answer” question.
- ❑ Precise computation rarely mentioned in publication.

What do trialists want to know?

Question of interest	Summary measure for one treatment group	Summary measure for treatment comparison and proportional hazards	Summary measure for treatment comparison and non-proportional hazards
Precision	KM confidence bands	Hazard ratio confidence interval	Hazard ratio confidence interval
Reliability	KM confidence bands, no. at risk	Assessment of proportional hazards assumption	-
Stability	Eg assume censored observations to be events	Assessment of proportional hazards assumption	-
Information		Information fraction	Depends on effect measure

Conclusion

- Follow-up quantifiers used in literature **highly heterogeneous**.
- Focus on **scientific question**, answering that using suitable quantities: precision, stability, information, assumptions for any quantity of interest.
- No hope that any of these questions can be answered with one single number, however defined.

PH vs NPH

- Assumption matters for stability.
- NPH: need to choose effect measure.
- Information depends on #events (PH) or many more quantities (NPH).

Resources

- Paper: https://arxiv.org/abs/2206.05216
- Oncology estimand WG: http://www.oncoestimand.org

Your Turn

- QUESTIONS?
  - The chair will moderate questions to the Panel

End

- THANK YOU!