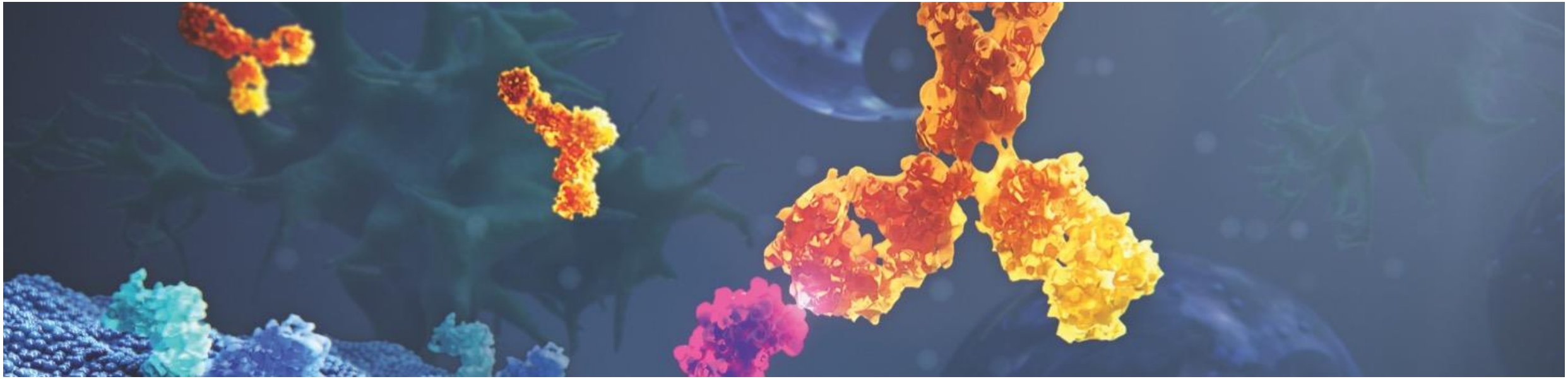
Causal Estimand and Principal Stratum, an Overview and Potential Applications in Oncology

Feng Liu, PhD on behalf of the Causal Subteam; Oncology Estimand Working Group

ASA Regulatory Industry Workshop 2019, September, 2019



Oncology Estimands WG

- Initiated and led by Evgeny Degtyarev (Novartis) and Kaspar Rufibach (Roche), first TC Feb 2018
- 32 members (14 from Europe and 18 from US) representing 20 companies

Estimands in Oncology:
Need for the Industry Working Group

- increased transparency on treatment effect of interest considered as important goal of the ICH E9 addendum

But **what if the same estimand is described differently by sponsors** in protocols and publications?

→ **confusion for HA, payers, physicians and patients**

→ **inconsistent labels**

→ more HA questions on estimands creating perception of estimand topic being rather a burden

- main **purpose of the Working Group**:
  - **ensure common understanding and consistent definitions for key estimands in Oncology across industry**
  - share experience and discuss estimands, intercurrent events and the used sensitivity analyses in Oncology

Why Causal estimand?

- ICHE9 addendum didn't explicitly state "causal" and "causal thinking" is made implicitly via referencing potential outcomes and adoption of the principal stratum strategy.
- Causal interpretations in oncology endpoint (Kaspar 2018)
  - Hazard ratio: depending on intercurrent events and causal thinking not clear (Kaspar 2018)
  - Average hazard ratio: Under PH or NPH assumption
  - Other endpoints (Landmark, RMST etc)

Causal-Subteam (Do we have an updated teamlist?)

- Kaspar Rufibach (Roche), lead
- Vera Beckers (Abbvie)
- Björn Bornkamp (Novartis)
- Audrey Boruvka (Roche)
- Andreas Brandt (BfARM)
- Marie-Laure Casadebaig (Celgene)
- Feng Liu (AstraZeneca)
- Yi Liu (Nektar)
- Juliane Manitz (EMD Serono)
- Emily Martin (EMD Serono)
- Devan Mehrotra (Merck)
- Alan Phillips (ICON)
- Satrajit Roychoudhury (Pfizer)
- Anja Schiel (NoMA)
- An Vandebosch (Janssen)

Agenda

- Clinical Questions
- Ideas behind Causal Estimand and Principle Stratification
- Estimation of Principal Stratum Effects
- Criticisms
- Summary

Introduction to Casual Inference

- Estimand
  - A population parameter that quantifies the effect of treatment relative to control.
  - Causally interpretable (NAS report)
- Definition of causality?
  - the process of drawing a conclusion about a causal connection based on the conditions of the occurrence of an effect.
- Causal inference: Does a relation from cause to effect exist?
- In the health sciences, many of the critical questions are causal in nature
- For example:
  - What is the efficacy of a given drug on a target population?
  - What fraction of HIV infections could have been prevented by a given treatment or policy?

Evaluation Question and Attributing Causality

What is the effect of an intervention/treatment P on outcome Y ?
Example: What is the effect of an intervention/treatment (P) on improvement in Overall Survival (Y)?

Impact of P =
OS (Y) for a cancer patient receiving intervention vs
OS (Y) for the same patients in the absence of the intervention
(at the same point in time)

We observe Y for cancer patients receiving intervention

But we do not observe Y for the same patient with receiving intervention.

Fundamental problem: We never observe the same individual with and without intervention at the same point in time

What if under post-treatment intercurrent events?

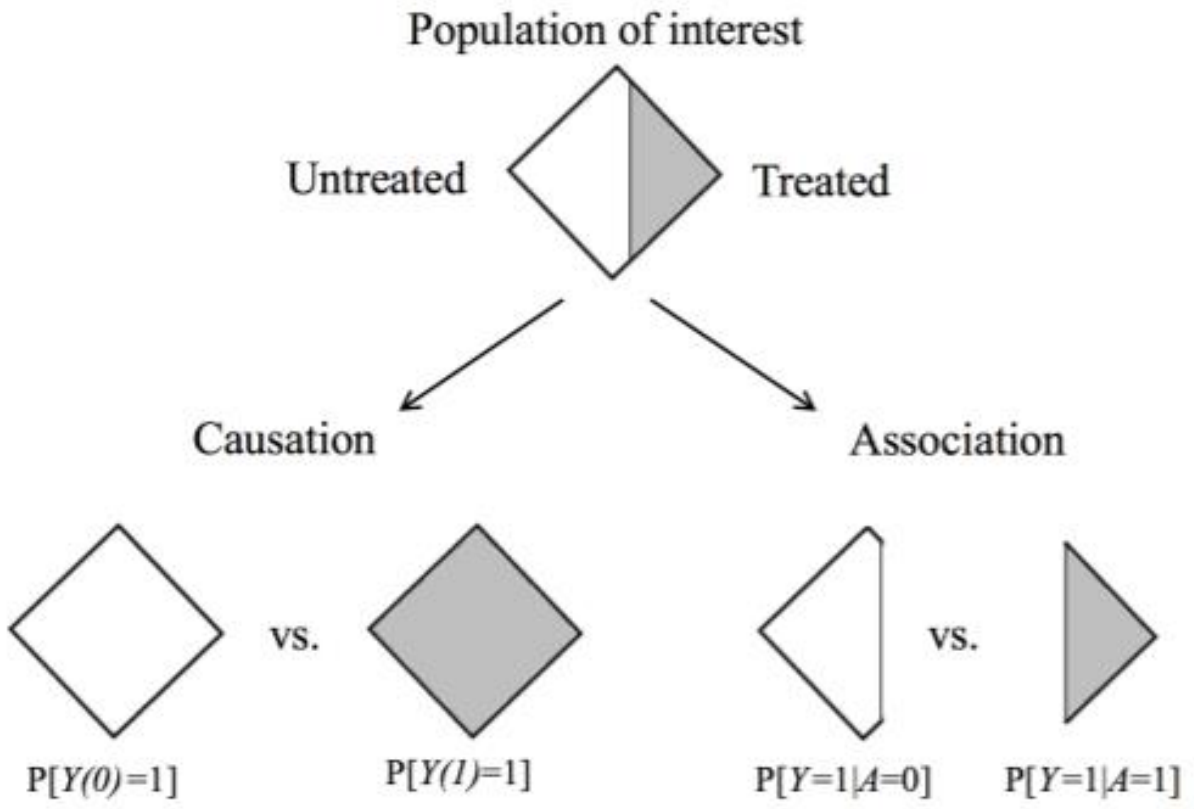
Attributing Causality

Estimate/ mimic/find a good proxy for what would have happened to outcome Y in the absence of program P

- Compare the patients with someone who ‘looks’ exactly like him/her who was not exposed to the intervention P at the same point of time

- In other words, we must find a valid Counterfactual or Control group

Identifying Causal impact Causation is not Correlation



Association: measures difference in risk between disjoint subsets of the population determined by individual's actual treatment value

Causation: measures difference in risk in the entire population under two treatment values

Evaluate the impact/effect of an intervention on some outcomes of interest

- By how much did X (intervention) change Y (outcome)?

Not the same as correlation!

- X and Y are related, move together in some way

Scharfstein 2017

Clinical Questions

- Antidrug antibodies (ADA)
  - For large molecular i.e. oncology immunotherapies: ADAs might form and may (or may not) have a neutralizing effect on the treatment
- What is the treatment effect versus control (e.g. on overall survival) in patients that develop ADAs if on the investigational treatment?
  - Note the control treatment might be a non-biologic drug, so that ADAs by definition will not form

Clinical Questions

- Commonality
  - Evaluate the treatment effect in the subgroup of patients where a specific post-randomization event would (or would not) occur
  - Challenge: Post-randomization event itself may be affected by treatment
    - Randomization cannot be relied upon to ensure comparable groups on investigational treatment and control → Selection bias
- Class of questions is quite frequent in oncology (effect in sub-population)
  - See EMA anticancer guidance (Section 7.6.5)on “Analyses based on a grouping of patients on an outcome of treatment”
    - Highlights problematic nature of naive analyses

Principal Stratification Estimands

- Concept introduced in Frangakis & Rubin (2002)
  - Introduce potential outcomes (binary) S(0) and S(1) for **every** patient in the trial
    - Even though just one of the two is observed for every patient
  - Determine treatment effect in subset(s) (principal strata) of population defined by S(0) and S(1) → leads to 4 principal strata

S – occurence of postbaseline event
S(0) – potential outcome control
S(1) – potential outcome treatment

- Example
  - Suppose we are interested in the treatment effect in patients, who develop ADAs on treatment (have S(1) = 1) (union of 2 principal strata)
  - For patients on treatment we observe S(1)
  - Problem: For patients on control we do not observe S(1)

Principal Stratification Estimands

- Provide a way to formulate the question/problem not the solution
- Why is this of any help then?
  - Provides a clear inferential target (treatment effect in principal strata)
  - Easier to discuss assumptions etc if inferential target is clear
- Determination of treatment effects in principal strata requires assumptions!
  - E.g. Principal stratum membership is not observed
- Let‘s illustrate with the ADA example in more detail

ADA example in more detail

- Quantity of interest?
  - Survival time under treatment or control for patients who would develop ADAs if given active treatment (S(1)=1).

Potential outcomes
Y(z) – Potential survival time
S(z) – ADA presence post-baseline

- In potential outcome notation: Compare

$$Y(1)|\{S(1) = 1\} \text{ versus } Y(0)|\{S(1) = 1\}$$

  - e.g. estimate survival functions $P(Y(1) > t|S(1) = 1)$ and $P(Y(0) > t|S(1) = 1)$ and derive a summary measure based on those

ADA example

- In potential outcome notation: Compare
  $$Y(1)|\{S(1) = 1\} \text{ versus } Y(0)|\{S(1) = 1\}$$
  - e.g. estimate survival functions $P(Y(1) > t|S(1) = 1)$ and $P(Y(0) > t|S(1) = 1)$ and derive a summary measure
- Easy to derive an estimate for $P(Y(1) > t|S(1) = 1)$: Observed on treatment arm
- How to derive estimate of $P(T(0) > t|S(1) = 1)$?
  - No one-size-fits-all solution in the Frangakis and Rubin (2002) paper

ADA example: Full Bayesian estimation

- We know that we observe a mixture of patients on the control arm

$$p(Y(0)) = \pi\, p(Y(0)|S(1) = 1) + (1 - \pi)\, p(Y(0)|S(1) = 0)$$

  - $\pi = P(S(1) = 1)$ can be estimated from the treatment arm
- Densities
  - $p(Y(0)|S(1) = 1)$ and $p(Y(0)|S(1) = 0)$ not identified based on the data without further (e.g. parametric) assumptions
- Binary outcome data
  - Even parametric assumptions not sufficient
  - Magnusson et al. (2018) utilize fully Bayesian approach for identification: Proper prior leads to a proper posterior distribution

    → Need to evaluate impact of “weakly-informative“ priors carefully

ADA example: Utilizing covariates

- Assume one can find all covariates X such that
  - Conditional on covariates X, Y(0) and S(1) are independent: $Y(0) \perp S(1) \mid X$
    - Principal ignorability, see Ding et al. 2017, Feller et al. 2017
    - Similar to assumptions used in propensity score matching analyse (Austin 2010, 2014)
  - Average treatment over population Estimand: Average treatment effect (ATE) E[Yi(1)-Yi(0)] ,
  - Sensitivity analysis: other unmeasured confounding covariates
  - If this is true the conditional distribution $p(Y(0) \mid S(1), X) = p(Y(0) \mid X)$

ADA example: Utilizing covariates

- Estimation (see also Bornkamp & Bermann, 2019)
  - Estimate p(Y(0) | X) on control group, average with respect to p(X | S(1) = 1) (regression adjustment/standardization)
  - Alternative estimation strategies
    - Multiple imputation of S(1) based on X
    - Matching on X and “ standard“ analysis

ADA example: Utilizing covariates

- Case-specific whether one would be willing to make this assumption
  - Principal ignorability: untestable assumption (independence assumption “across worlds“); sensitivity analyses possible, see Ding et al. (2017)
  - If S(0) would be predictive of S(1) further analyses/assumptions would be possible → in this case as S(0) = 0 for all patients

Criticisms

- Complication: Benefit-risk analyses for principal strata
  - Typical analysis strategies do not clearly identify the population of patients in the principal stratum. How to perform safety analyses?
- Hernán & Scharfstein (2018)
  - "... subgroup that cannot be clinically identified ..."
- Scharfstein (2018)
  - Principal stratification is scientifically interesting but just too assumption-laden to be primary
  - "... Lowers the level of evidence. ..."

Estimands (Scharfstein 2017)

Estimands	Causal thinking in Hypothesis Testing	Assumptions
Treatment Policy	P[Y(1)=1] vs P[Y(0)=1]	ignore intercurrent events
Hypothetical	P[Y(1,R(0))=1] vs P[Y(0, R(0))=1]	
Composite	P[U(1)=1] vs P[U(0)=1] U conditioned on outcome Y and intercurrent events	ITT effect on composite outcome (outcome Y and IE)
Principle Stratum	P[Y(1)=1|R(1)=0,R(0)=0] vs P[Y(0)=1|R(1)=0,R(0)=0]	

Summary

- ICHE9 addendum added causal thinking i.e. principal stratum
- Casual thinking is nature in oncology with added complexity of intercurrent events
  - Treatment policy vs principal stratum in handling intercurrent events
  - Sometimes assumptions considered too strong to answer questions
  - utilizing causal inference techniques will raise the level of discussion on the questions and possible assumptions which leads to
  - More work needed on: What are plausible assumptions (& thus analyses)?
- Ways to estimate “average causal effect”
- Due to assumptions required for identification, the principal stratum strategy might not be part of the primary estimand
- Important to contribute to an “overall” picture of the drug’s properties

References

- Bornkamp, B. and Bermann, G. (2019) Estimating the treatment effect in a subgroup defined by an early post-baseline biomarker measurement in randomized clinical trials with time-to-event endpoint. Statistics in Biopharmaceutical Research (to appear).
- Ding, P. and Lu, J. (2017) Principal stratification analysis using principal scores." Journal of the Royal Statistical Society: Series B 79: 757-777.
- Feller, A., Mealli, F. and Miratrix, L. (2017) Principal score methods: Assumptions, extensions, and practical considerations. Journal of Educational and Behavioral Statistics 42: 726-758.
- Frangakis, C. E. and Rubin, D. (2002) Principal stratification in causal inference. Biometrics 58: 21-29.
- Hernán, M. A. and Scharfstein, D. (2018) Cautions as Regulators Move to End Exclusive Reliance on Intention to Treat. Annals of internal medicine 168: 515-516.
- Magnusson, B., Schmidli, H., Rouyrre, N. and Scharfstein, D. (2018) Bayesian inference for a principal stratum estimand to assess the treatment effect in a subgroup characterized by post-randomization events, arXiv:1809.03741
- Pearl, J. (2011) Principal stratification -- a goal or a tool? The International Journal of Biostatistics 7: 1-13.
- Scharfstein, D. (2018) A (Constructive/Provocative) Critique of the ICH E9 Addendum, Presentation given at 11th Annual Conference on Statistical Issues in Clinical Trials Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania https://www.cceb.med.upenn.edu/sites/default/files/uploads/cci/2018_Clin_Trials/ScharfsteinD%20materials.pdf
- Kaspar 2018 Treatment effect quantification for time-to-event endpoints–Estimands, analysis strategies, and beyond