ICH E9 addendum and principal stratification

Kaspar Rufibach
Methods, Collaboration, and Outreach Group, Roche Data and Statistical Sciences
DIA China
24th August 2021

Acknowledgments

This material:

- was first presented at the 76th Deming Conference on Applied Statistics on 9th December 2020,
- by Kaspar Rufibach and **Evgeny Degtyarev (Novartis)**.
- Section on CAR-T initially prepared by Evgeny.

Acknowledgments

All our colleagues of the **industry working group on estimands in oncology**.

Regulatory colleagues around the world for regular discussion, their input, and feedback.

Agenda

1. Subgroups by post-randomization event - principal stratification
2. Case study: CAR-T
3. Backup: Industry working group *Estimands in oncology*
4. Backup: Estimation of average causal effect
5. Backup: Estimation of principal effects

Agenda

1. Subgroups by post-randomization event - principal stratification
2. Case study: CAR-T
3. Backup: Industry working group *Estimands in oncology*
4. Backup: Estimation of average causal effect
5. Backup: Estimation of principal effects

Subgroups by post-randomization event - principal stratification

"... The target population might be taken to be the "principal stratum" in which an **intercurrent event would occur**. Alternatively, the target population might be taken to be the principal stratum in which an intercurrent event **would not occur**. The clinical question of interest relates to the treatment effect only within the principal stratum..."

ICH E9 working group (2019)

Principal stratification:

- Originates in causal inference: Frangakis and Rubin (2002).
- Framework for comparing treatments adjusting for **posttreatment** variables.
- Formulated within **potential outcomes** framework.
- Yields principal effects which are **causal** effects within a principal stratum.

Introductory books causal inference: Imbens and Rubin (2015), Hernán and Robins (2020).

First, let us summarize what does not work.

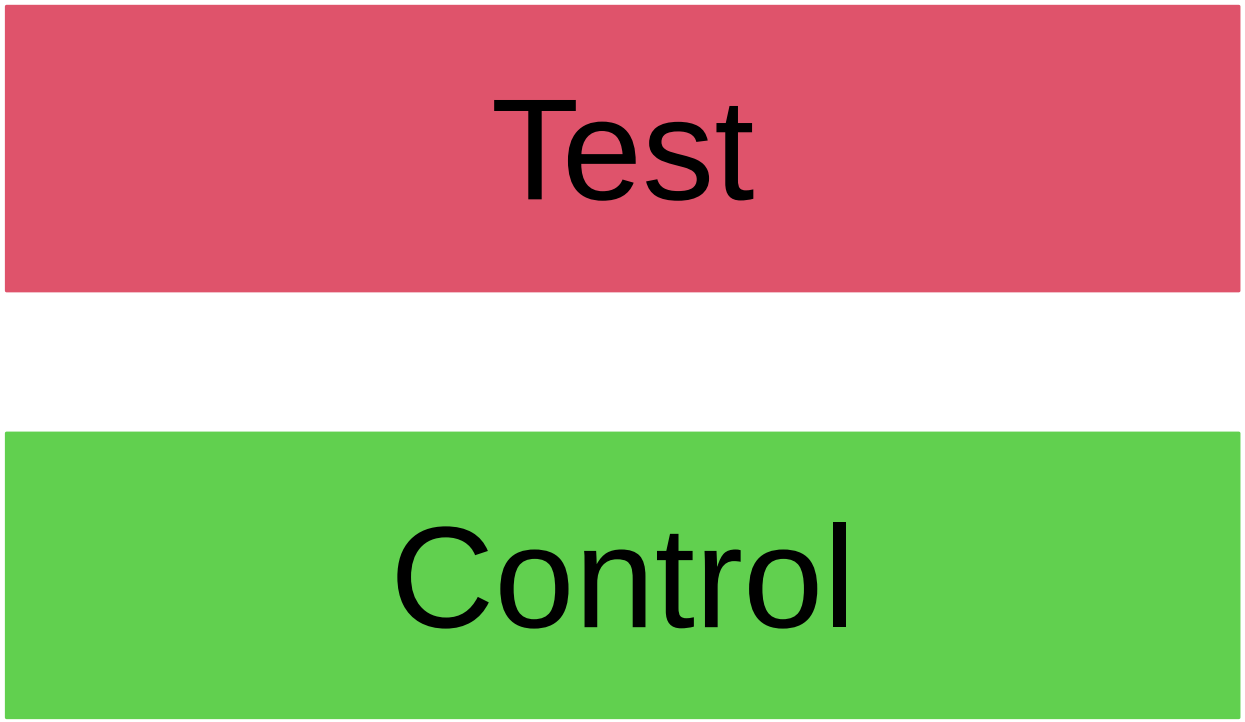
2-arm RCT test (T) vs. control (C)

Do responders
have higher treatment effect?

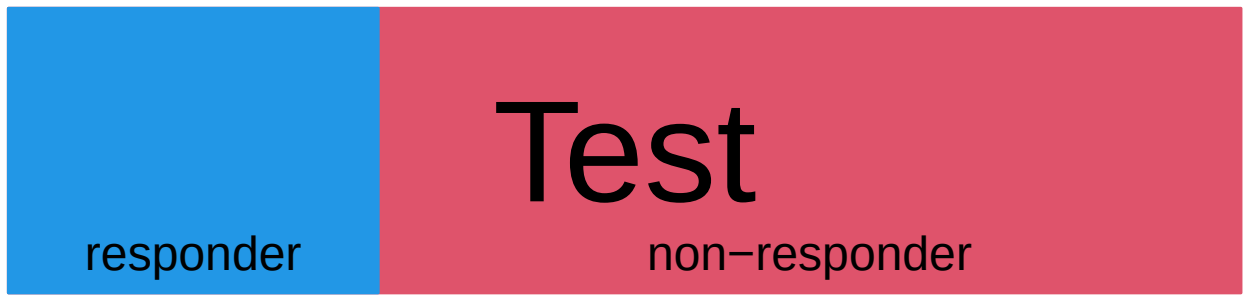
"Subgroup" built by post-randomization event!

How can we make valid causal statements?

Need "matched control patients"!


Test
Control


Test
responder
non−responder



Patients who respond
if randomized to Test
had they received control

Test

Control

Test

Control


Test
responder
non−responder

For every complex problem, there is a solution that is simple, neat, and wrong.

H.L. Mencken, American Journalist

Naive analyses are misleading and do not answer causal question

Principal stratification: "subgroup analysis for post-baseline subgroups"

randomization + assumptions

Are such questions relevant?

TABLE 2 Examples of principal stratum estimands discussed in this section

Example	Scientific question	Primary endpoint	Intercurrent event	Stratum of interest
Multiple sclerosis	Treatment effect on confirmed disability progression in the subpopulation of relapse-free patients	Time to confirmed disability progression	Post-randomization relapse	Patients who would be relapse-free under both treatments
Treatment effect in early responders	Predict treatment effect on long-term primary endpoint based on early biomarker-type readout	Time-to-event	Biomarker value above or below a pre-specified threshold	Patients who would respond early under treatment vs. those that would not
Antidrug antibodies (ADA) for targeted oncology drugs	Do patients that develop ADAs on either arm still benefit from the drug?	Time-to-event	Development of antidrug antibodies because of receiving treatment	Patients who would be ADA+ under treatment
Impact of exposure on OS	Do patients with insufficient exposure have lower treatment effect?	Time-to-event	Exposure below a pre-specified threshold	Patients with low vs. non-low exposure under treatment
Prostate cancer prevention	Assess effect of treatment to prevent prostate cancer on severity of prostate cancer among those men who would be diagnosed with prostate cancer regardless of their treatment assignment	Time-to-event	Getting prostate cancer	Patients who get prostate cancer irrespective of treatment

Bornkamp et al. (2021).

CAR-T example - see later!

OS / PFS by response.

Potential outcomes and principal stratification

$$Z := \begin{cases} 1 & \text{test treatment} \\ 0 & \text{control treatment.} \end{cases}$$

Y: outcome (binary, continuous, time-to-event).

Ideal world: treating physician decides on treatment based on outcome if given

- control treatment: $Y(Z = 0) = Y(0)$,
- test treatment, $Y(Z = 1) = Y(1)$.

Neither $Y(0)$ nor $Y(1)$ known when assigning treatment!

Only one observed at all $\Rightarrow$ **individual causal effect** $Y(1) - Y(0)$ not observed.

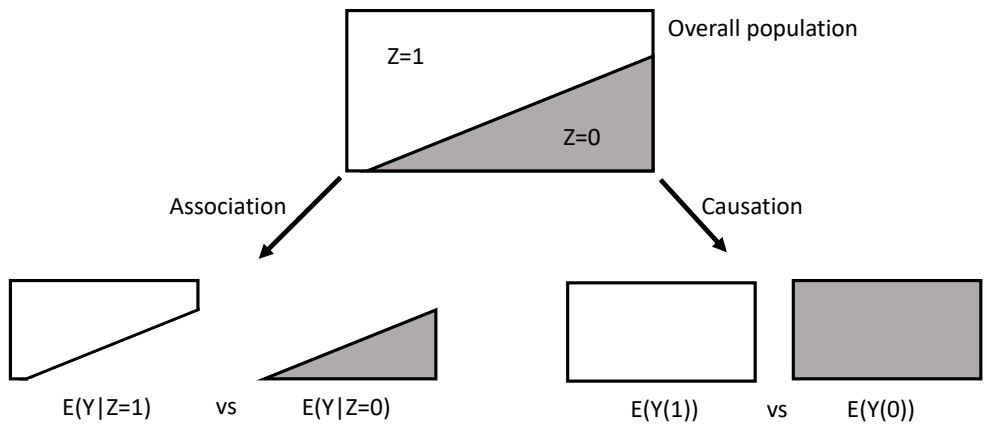
What are causal effects?

$Y(1)_i$: potential outcome for patient i.

$\mathcal{S}$: population of patients.

Causal treatment effect:

- Comparison of $\{Y(1)_i, i \in \mathcal{S}\}$ vs. $\{Y(0)_i, i \in \mathcal{S}\}$.
- Compare outcomes "had everyone received treatment" vs. outcomes "had everyone received control". **Hypothetical scenario**.



Naive analysis

Not a causal effect: comparison of $\{Y(1)_i, i \in \mathcal{S}_1\}$ vs. $\{Y(0)_i, i \in \mathcal{S}_2\}$ with $\mathcal{S}_1 \neq \mathcal{S}_2$.

Naive analysis: Let S = indicator variable for intercurrent event, e.g. responder.

- Compare patients with $S = 1$ on both test and control arm.
- RCT: $S(Z)$ **post-randomization** $\Rightarrow$ S **depends on** Z!
- We observe $S(Z = 1)$ on test and $S(Z = 0)$ on control $\Rightarrow$ population of patients with $S(1) = 1$ and $S(0) = 1$ might be **quite different**!
- Breaks randomization $\Rightarrow$ not comparing "like with like" $\Rightarrow$ **not estimating causal effect**.
- Numerically observe a treatment effect in naive analysis $\Rightarrow$ not clear whether
  - due to different treatments or
  - due to difference in compared populations.
- Estimates treatment effect in principal stratum $\{S(1) = 1\} \cap \{S(0) = 1\}$ assuming $S(1) = S(0) \Rightarrow$ response not treatment related. Assumption quite strong and **rarely justified**!

Principal stratification

Idea: stratify patients based on **potential outcomes** $S(0), S(1)$ for **all** treatments.

	$S(0) = 1$	$S(0) = 0$
$S(1) = 1$	$\{S(1) = 1\} \cap \{S(0) = 1\}$	$\{S(1) = 1\} \cap \{S(0) = 0\}$
$S(1) = 0$	$\{S(1) = 0\} \cap \{S(0) = 1\}$	$\{S(1) = 0\} \cap \{S(0) = 0\}$

Causal interpretation:

- Stratify population according to the **same rule on treatment and control arm**.
- Possible since membership to principal stratum **fixed at baseline**, not affected by treatment assignment.

Caveat:

- For patients on test arm we observe $S(1)$, but not $S(0)$, and vice versa for patients on control arm.
- **Identification** of patients in strata of interest generally not possible, not even after observing Y and S in a given trial.

Example: antidrug antibodies in immunotherapies

- Biological drugs: may trigger immune responses $\Rightarrow$ formation of **antidrug antibodies** (ADAs).
- Scientific question: Do patients that develop ADAs still benefit from the drug?
- Y: PFS or OS.
- S: occurrence of ADA at x weeks, say $x = 4$.
- Depending on test and control treatment $\Rightarrow$ ADA only in test arm.

	$S(0) = 1$	$S(0) = 0$
$S(1) = 1$	$\{S(1) = 1\} \cap \{S(0) = 1\}$	$\{S(1) = 1\} \cap \{S(0) = 0\}$
$S(1) = 0$	$\{S(1) = 0\} \cap \{S(0) = 1\}$	$\{S(1) = 0\} \cap \{S(0) = 0\}$

	ADA+ under control	ADA- under control
ADA+ under test	**Stratum of interest**	
ADA- under test		

Effect measures

Primary interest:

- Compare $Y(1)$ vs. $Y(0)$ in stratum $\{S(1) = 1\}$.
- Contrast this to results in $\{S(1) = 0\}$.

Effect measure:

- (Hazard ratio **not causally interpretable**: Aalen et al. (2015).)
- Base effect measure on **survival functions**:

$$U_1(t) := P(Y(1) > t | S(1) = 1) \quad \text{and} \quad U_0(t) := P(Y(0) > t | S(1) = 1).$$

Examples:

- **Milestone** difference at $t^* > \tilde{t}$:

$$\delta(t^*) = U_1(t^*) - U_0(t^*).$$

- Time-averaged version, i.e. difference in **RMST**:

$$\int_0^{t^*} \delta(t)dt = E[\min(Y(1), t^*) - \min(Y(0), t^*)].$$

Potential outcomes, estimands, and PS

All estimand strategies can be formulated using potential outcomes:
Lipkovich et al. (2020).

Additional complications: Y time-to-event $\Rightarrow$ outcome event = competing risk for intercurrent event. Naive analyses conditioning on observed intercurrent event:

- Compares **non-randomized** populations.
- **Immortal bias**: patients immortal until observation of S.

Sensitivity analyses!

Assumptions for estimation (see backup) **unverifiable**:

- "Across-world" $\Rightarrow$ even with **infinite number of observations** we could not test them.
- Only verifiable if we could observe both, patient receives control in one world and treatment in other.

scientific knowledge + sensitivity analyses

Conclusions principal stratification

Conclusions:

- Many relevant examples in drug development.
- Scientific question typically not primary, but important to characterize treatment effect in subgroups built by intercurrent events, such as ADA or CAR-T. **Both explicitly requested by HAs**!
- Naive analyses often standard: Unclear estimand $\Rightarrow$ **causal conclusion unclear**.
- Complex question $\Rightarrow$ complex analysis needed.
- Assumptions needed: scientific input + sensitivity analyses.

MAIN PAPER

Principal stratum strategy: Potential role in drug development

Björn Bornkamp[1] | Kaspar Rufibach[2] | Jianchang Lin[3] | Yi Liu[4] | Devan V. Mehrotra[5] | Satrajit Roychoudhury[6] | Heinz Schmidli[1] | Yue Shentu[7] | Marcel Wolbers[2]

Bornkamp et al. (2021)

Markdown:

https://oncoestimand.github.io/princ_strat_drug_dev/princ_strat_example.html

BBS seminar:

http://bbs.ceb-institute.org/?p=1587

Effective statistician podcast, Björn Bornkamp and Kaspar Rufibach:

https://theeffectivestatistician.com/

a-deep-dive-into-principal-stratification-and-causal-inference

arXiv.org > stat > arXiv:2101.04263

Search...

Help | Ad

Statistics > Methodology

[Submitted on 12 Jan 2021]

Weighted Approach for Estimating Effects in Principal Strata with Missing Data for a Categorical Post-Baseline Variable in Randomized Controlled Trials

Shengchun Kong, Dominik Heinzmann, Sabine Lauer, Tian Lu

This research was motivated by studying anti-drug antibody (ADA) formation and its potential impact on long-term benefit of a biologic treatment in a randomized controlled trial, in which ADA status was not only unobserved in the control arm but also in a subset of patients from the experimental treatment arm. Recent literature considers the principal stratum estimand strategy to estimate treatment effect in groups of patients defined by an intercurrent status, i.e. in groups defined by a post-randomization variable only observed in one arm and potentially associated with the outcome. However, status information might be missing even for a non-negligible number of patients in the experimental arm. For this setting, a novel weighted principal stratum approach is presented: Data from patients with missing intercurrent event status were re-weighted based on baseline covariates and additional longitudinal information. A theoretical justification of the proposed approach is provided for different types of outcomes, and assumptions allowing for causal conclusions on treatment effect are specified and investigated. Simulations demonstrated that the proposed method yielded valid inference and was robust against certain violations of assumptions. The method was shown to perform well in a clinical study with ADA status as an intercurrent event.

Kong et al. (2021)

Github repository: https://github.com/openpharma/BBS-causality-training

Talk Dominik in BBS seminar: http://bbs.ceb-institute.org/?p=1668

Design trumps analysis.

Don Rubin, American Statistician

Rubin (2008)

Agenda

1. Subgroups by post-randomization event - principal stratification
2. Case study: CAR-T
3. Backup: Industry working group *Estimands in oncology*
4. Backup: Estimation of average causal effect
5. Backup: Estimation of principal effects

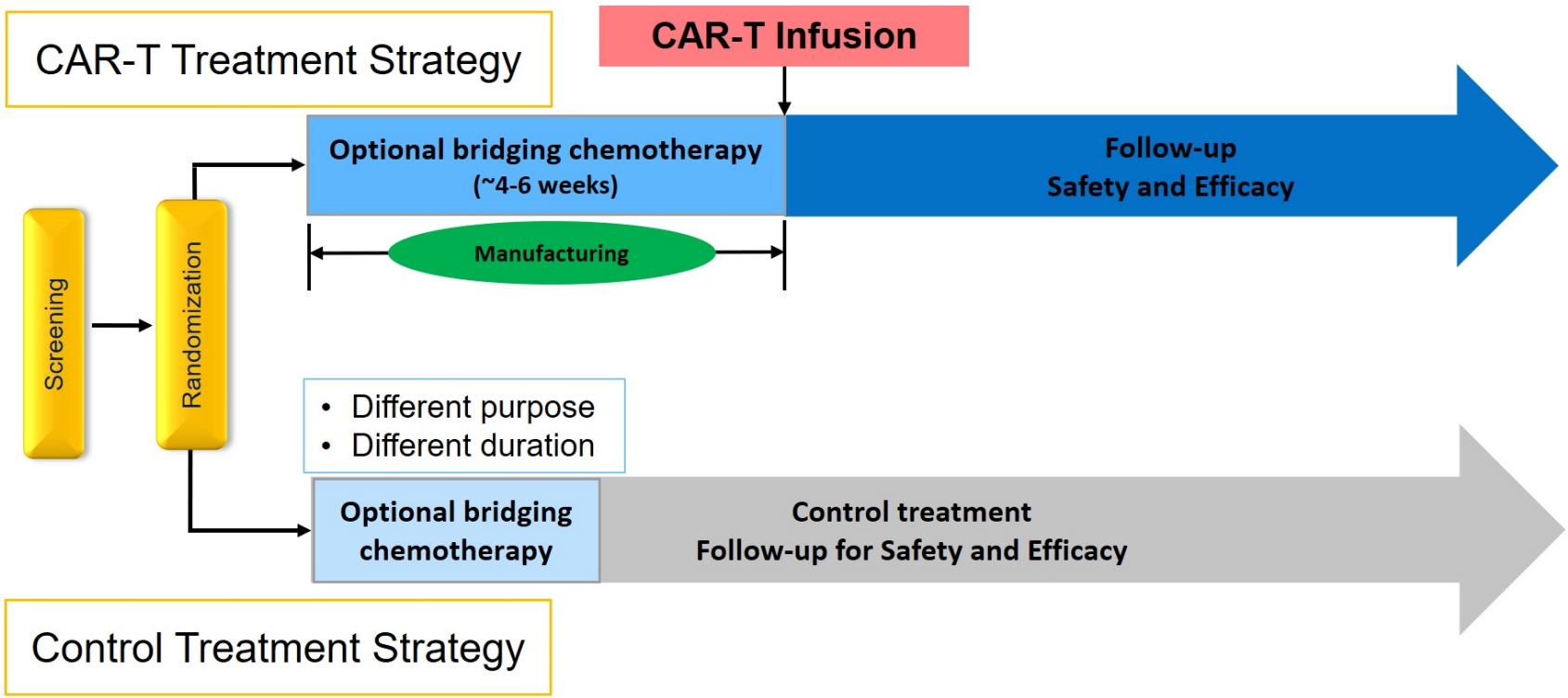
Case study: CAR-T

CAR-T example - acknowledgments

Based on example:

- Presented by Aiesha Zia
- at BBS Webinar **RCTs meeting causal inference: principal stratum strategy and beyond** on September 7th, 2020.
- `http://bbs.ceb-institute.org/?p=1587`

RCT comparing two strategies



Primary scientific question: OS comparison of entire sequence of interventions.

FDA comment on the protocol

> *Subjects in the CAR-T arm may receive extensive bridging chemotherapy while awaiting CAR-T manufacture, and some, especially those experiencing extended delays in product manufacture, could achieve a CR/CRi [...] status in response to aggressive bridging chemotherapy even before initiation of CAR-T treatment. Since these responses cannot be directly attributed to CAR-T treatment, the statistical assessment plan should prospectively create rules for appropriately censoring CR [...] subjects.*

FDA's interest: effect for patients who do not respond to bridging in CAR-T.

Censoring?

FDA proposal for supplementary EFS analysis:

- **Censor** patients who respond to bridging chemotherapy in CAR-T arm.
- Censoring targets **hypothetical scenario** in which no patient would respond to bridging chemotherapy in CAR-T arm.
- Is this estimand relevant for patients, physicians, and regulators?

Getting the question right!

Analysis requested by FDA does not address relevant question of interest.

Sponsor: suggested **principal stratum estimand** would address FDA's actual question of interest.

Estimand:

- **Treatment**: CAR-T relative to control treatment strategy.
- **Variable**: EFS.
- **Intercurrent event**: none left.
- **Summary measure**: hazard ratio.
- **Population**:
  - **Principal stratification**: Effect in patients who would not respond to bridging chemotherapy if they were given bridging chemotherapy? Estimation: see backup.
  - **Hypothetical**: Effect if no patient would respond to bridging chemotherapy in CAR-T arm? Estimation: through censoring.

FDA: agreed to use principal stratum estimand as supplementary instead of hypothetical.

Patients randomized to
CAR−T not responding
to bridging
had they received control

Naive comparison

Naive comparison of observed non-responders:

- Only valid if all patients in control arm share same EFS regardless of their response to bridging chemotherapy if they were given bridging chemotherapy: $Y(Z = 0)$ and $S(1)$ independent.
- "Valid": unbiased estimate of causal treatment effect.

Principal stratum estimand

Opportunities:

- Discuss questions of interest and not censoring rules $\Rightarrow$ improved Health Authority interactions.
- More meaningful and interpretable analyses.

If you do not know how to ask the right question, you discover nothing.

W.E. Deming, American Statistician

Thank you for your attention.

kaspar.rufibach@roche.com
http://go.roche.com/dss-mco

http://www.kasparrufibach.ch
numbersman77
numbersman77

References I

- Aalen, O. O., Cook, R. J. and Røysland, K. (2015). Does Cox analysis of a randomized survival study yield a causal treatment effect? *Lifetime Data Anal* **21** 579–593.
- Angrist, J. D., Imbens, G. W. and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* **91** 444–455.
  http://www.jstor.org/stable/2291629
- Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res* **46** 399–424.
- Bornkamp, B., Rufibach, K., Lin, J., Liu, Y., Mehrotra, D. V., Roychoudhury, S., Schmidli, H., Shentu, Y. and Wolbers, M. (2021). Principal stratum strategy: Potential role in drug development. *Pharm Stat, to appear* .
- Casey, M., Degtyarev, E., Lechuga, M. J., Aimone, P., Ravaud, A., Motzer, R. J., Liu, F., Stalbovskaya, V., Tang, R., Butler, E., Sailer, O., Halabi, S. and George, D. (2021). Estimand framework: Are we asking the right questions? a case study in the solid tumor setting. *Pharmaceutical Statistics* **20** 324–334.
- Degtyarev, E., Rufibach, K., Shentu, Y., Yung, G., Casey, M., Englert, S., Liu, F., Liu, Y., Sailer, O., Siegel, J., Sun, S., Tang, R., Zhou, J. and on behalf of the Industry Working Group on Estimands in Oncology (2020). Assessing the impact of covid-19 on the clinical trial objective and analysis of oncology clinical trials - application of the estimand framework. *Statistics in Biopharmaceutical Research* **12** 427–437.
  https://doi.org/10.1080/19466315.2020.1785543
- European Medicines Agency, Committee for Medicinal Products for Human Use (2019). Mayzent: Assessment report.
  https://www.ema.europa.eu/en/documents/assessment-report/mayzent-epar-public-assessment-report_en.pdf
- Frangakis, C. E. and Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics* **58** 21–29.
- Hernán, M. A. and Robins, J. M. (2020). *Causal Inference: What If*. Chapman & Hall/CRC.
  https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/
- ICH E9 working group (2019). ICH E9 (R1): addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials.

References II

- Imbens, G. W. and Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.
- Joffe, M. M., Small, D., Hsu, C.-Y. et al. (2007). Defining and estimating intervention effects for groups that will develop an auxiliary outcome. *Statistical Science* **22** 74–97.
- Kong, S., Heinzmann, D., Lauer, S. and Lu, T. (2021). Weighted approach for estimating effects in principal strata with missing data for a categorical post-baseline variable in randomized controlled trials. `https://arxiv.org/abs/2101.04263`
- Lawrance, R., Degtyarev, E., Griffiths, P., Trask, P., Lau, H., D'Alessio, D., Griebsch, I., Wallenstein, G., Cocks, K. and Rufibach, K. (2020). What is an estimand & how does it relate to quantifying the effect of treatment on patient-reported quality of life outcomes in clinical trials? *J Patient Rep Outcomes* **4** 68.
- Lipkovich, I., Ratitch, B. and Mallinckrodt, C. H. (2020). Causal inference and estimands in clinical trials. *Statistics in Biopharmaceutical Research* **1** 54–67.
- Magnusson, B. P., Schmidli, H., Rouyrre, N. and Scharfstein, D. O. (2019). Bayesian inference for a principal stratum estimand to assess the treatment effect in a subgroup characterized by postrandomization event occurrence. *Statistics in Medicine* **38** 4761–4771. `https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.8333`
- Manitz, J., Kan-Dobrosky, N., Buchner, H., Casadebaig, M., Degtyarev, E., Dey, J., Haddad, V., Fei, J., Martin, E., Mo, M., Rufibach, K., Shentu, Y., Stalbovskaya, V., Tang, R., Yung, G. and Zhu, J. (2021). Estimands in clinical trials with treatment switching. *Pharm Stat, to apper* .
- Pearl, J. (2009). *Reflections, Elaborations, and Discussions with Readers*. Cambridge University Press, 331–400.
- Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics* **2** 808–840.
- Sun, S., Weber, H.-J., Butler, E., Rufibach, K. and Roychoudhury, S. (2021). Estimands in hematologic oncology trials. *Pharmaceutical Statistics* **20** 793–805.

Backup

Agenda

1. Subgroups by post-randomization event - principal stratification
2. Case study: CAR-T
3. Backup: Industry working group *Estimands in oncology*
4. Backup: Estimation of average causal effect
5. Backup: Estimation of principal effects

Backup: Industry working group *Estimands in oncology*

Industry working group on estimands in oncology:

- Founded February 2018.
- Represents industry in Europe and US:
  - European special interest group "Estimands in oncology", sponsored by PSI and EFSPI.
  - ASA scientific working group of ASA biopharmaceutical section.
- **68** members (30 EU + 31 US + 7 Asia) representing **35** companies.
- Regularly interacts with **8 health authorities**.
- Presentations, webinars, papers.

www.oncoestimand.org


NOVARTIS
Roche
Pfizer
MERCK
janssen
PHARMACEUTICAL COMPANIES OF Johnson & Johnson
statisticon
SANOFI
gsk
abbvie
AstraZeneca
BeiGene
IQVIA
BAYER
Takeda
Celgene
astellas
SERVIER
modus OUTCOMES
MERCK
PPD
NEKTAR
Merus battling cancer
CLINICAL OUTCOMES SOLUTIONS
Innovative science to capture the patient voice
statburo Statistical Consulting
AMGEN
ICON
A Symbol of Excellence
Boehringer Ingelheim
phastar
Adelphi
ADELPHI VALUES
Daiichi-Sankyo
MERCK SERONO
zai Lab
Evidera
CANCER RESEARCH UK
BIRMINGHAM CANCER RESEARCH UK CLINICAL TRIALS UNIT
UNIVERSITY OF BIRMINGHAM

Papers

Published or accepted:

- Lawrance et al. (2020): What is an estimand & how does it relate to quantifying the effect of treatment on patient-reported quality of life outcomes in clinical trials. **link**
- Degtyarev et al. (2020): Assessing the impact of COVID-19 on the objective and analysis of oncology clinical trials - application of the estimand framework. **link**
- Casey et al. (2021): Estimand framework: Are we asking the right question? A case study in the solid tumor setting. **link**
- Sun et al. (2021): Estimands in Hematology Trials. **link**
- Manitz et al. (2021): Estimands in clinical trials with treatment switching.
- Bornkamp et al. (2021): Principal Stratum Strategy: Potential Role in Drug Development. **link** (incl. markdown).

More papers under preparation.

Current task forces

- Clinical engagement.
- *Principal stratification and treatment switching.*
- Time to response and DOR.
- Estimands and PRO.
- Follow-up quantification.
- RWD.
- Conditional vs. marginal.
- Time to event endpoints with prognostic or predictive biomarker subgroups.

Agenda

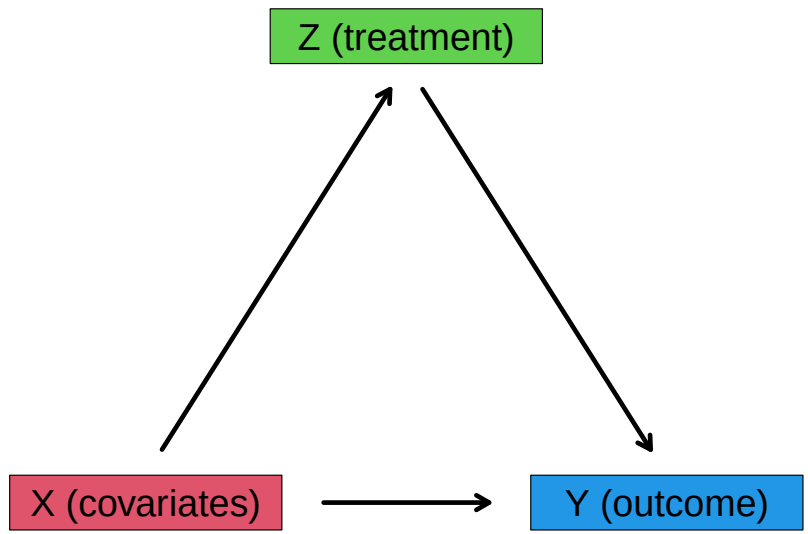
1. Subgroups by post-randomization event - principal stratification
2. Case study: CAR-T
3. Backup: Industry working group *Estimands in oncology*
4. Backup: Estimation of average causal effect
5. Backup: Estimation of principal effects

Backup: Estimation of average causal effect

Estimation of average causal effect

Key assumptions:

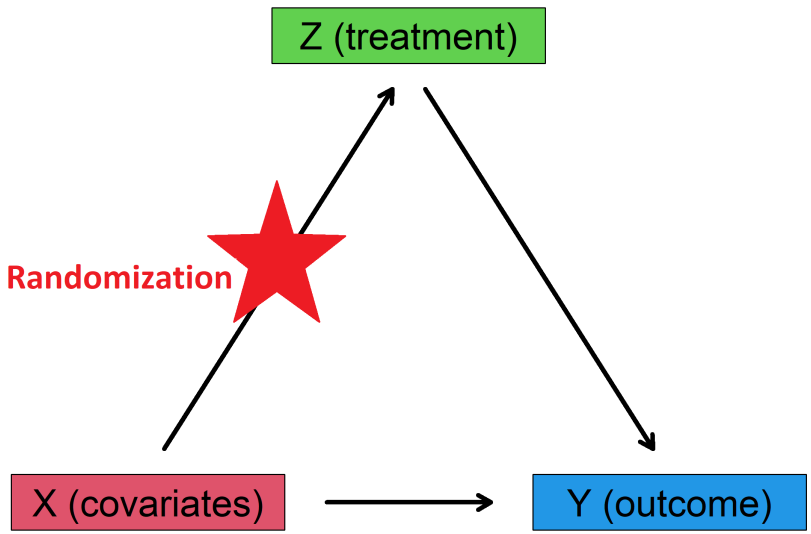
- **Exchangeability**: Counterfactual outcomes independent of treatment assignment $\Leftrightarrow$ $Y(1)$ and $Y(0)$ independent of Z. Trivially fulfilled in **RCT**. Via propensity scores otherwise.
- **Consistency**: No multiple versions of treatment $\Leftrightarrow$ individual's PO under observed exposure IS her observed outcome $\Leftrightarrow$ $E(Y(x)|Z = x) = E(Y|Z = x), x = 0, 1.$



Estimation of average causal effect

Key assumptions:

- **Exchangeability**: Counterfactual outcomes independent of treatment assignment $\Leftrightarrow$ $Y(1)$ and $Y(0)$ independent of Z. Trivially fulfilled in **RCT**. Via propensity scores otherwise.
- **Consistency**: No multiple versions of treatment $\Leftrightarrow$ individual's PO under observed exposure IS her observed outcome $\Leftrightarrow$ $E(Y(x)|Z = x) = E(Y|Z = x), x = 0, 1$.



Estimation of average causal effect

Key assumptions:

- **Exchangeability**: Counterfactual outcomes independent of treatment assignment $\Leftrightarrow$ $Y(1)$ and $Y(0)$ independent of Z. Trivially fulfilled in **RCT**. Via propensity scores otherwise.
- **Consistency**: No multiple versions of treatment $\Leftrightarrow$ individual's PO under observed exposure IS her observed outcome $\Leftrightarrow$ $E(Y(x)|Z = x) = E(Y|Z = x), x = 0, 1$.

$$
\begin{aligned}
E(Y(1) - Y(0)) &\overset{\textit{linearity of } E}{=} E(Y(1)) - E(Y(0)) \\
&\overset{\textit{exchangeability}}{=} E(Y(1)|Z = 1) - E(Y(0)|Z = 0) \\
&\overset{\textit{consistency}}{=} E(Y|Z = 1) - E(Y|Z = 0).
\end{aligned}
$$

So - why do we randomize?

- To balance covariates? **NO!**
- Covariates do not appear at all in above computation!
- Randomization generates equal distributions (in both groups) of **potential outcomes**!

For example, one would be extremely hard pressed to find a statistics textbook, even at the graduate level, containing a mathematical proof that randomization indeed produces unbiased estimates of the quantities we wish estimated – i.e., efficacy of treatments or policies.

Judea Pearl, American computer scientist and philosopher

Pearl (2009)

Estimation of average causal effect

Observational study:

- Decision between $Z = 0$ and $Z = 1$ might depend on X (measured or unmeasured).
- $Y(1)$ and $Y(0)$ **not** independent of $Z \Rightarrow$ exchangeability violated $\Rightarrow E(Y(1)) \neq E(Y(1)|Z = 1)$ and $E(Y(0)) \neq E(Y(0)|Z = 0)$.
- Patients who receive $Z = 1$ (for whom we observe $Y(1)$) might be **systematically different** from those who receive $Z = 0$ (for whom we observe $Y(0)$).
- Patients receiving $Z = 0$ **not representative** of overall population.

$$
\begin{aligned}
E(Y(1) - Y(0)) &\overset{\textit{linearity of } E}{=} E(Y(1)) - E(Y(0)) \\
&\overset{\cancel{\textit{exchangeability}}}{\neq} E(Y(1)|Z = 1) - E(Y(0)|Z = 0) \\
&\overset{\textit{consistency}}{=} E(Y|Z = 1) - E(Y|Z = 0).
\end{aligned}
$$

Agenda

1. Subgroups by post-randomization event - principal stratification
2. Case study: CAR-T
3. Backup: Industry working group *Estimands in oncology*
4. Backup: Estimation of average causal effect
5. Backup: Estimation of principal effects

Backup: Estimation of principal effects

Assumptions

Randomization not enough to estimate principal effects.

Need assumptions.

Estimation

SUTVA:

- Underpins virtually all estimation methods.
- **No interference** (treatment received by one individual does not affect PO of other individuals) + **no multiple versions** of treatment.
  - Infectious diseases: treatment may change depending on who else is vaccinated $\Rightarrow$ violation.

Monotonicity:

- $S(1) \geq S(0) \Rightarrow$ patients that are ADA+ on control would also be ADA+ on test.
- Patient with $S(0) = 1$ observed $\Rightarrow$ would know that $S(1) = 1 \Rightarrow$ bottom-left stratum in table empty.
- Allows estimation of principal stratum prevalences.

Estimation

Exclusion restriction:

- Assume $Y(0) = Y(1)$ (no treatment effect) for patients $\{S(0) = 0\} \cap \{S(1) = 0\}$ and $\{S(0) = 1\} \cap \{S(1) = 1\}$.

	$S(0) = 1$	$S(0) = 0$
$S(1) = 1$	no causal effect of Z on Y	$\{S(1) = 1\} \cap \{S(0) = 0\}$
$S(1) = 0$	$\{S(1) = 0\} \cap \{S(0) = 1\}$	no causal effect of Z on Y

- Treatment assignment Z (randomization in RCT) exclusively affects outcome through intercurrent event S.
- Angrist et al. (1996), Joffe et al. (2007).

Estimation

Joint models, Frangakis and Rubin (2002):

- Model for outcome given PS membership: $Y(0), Y(1)|S(1), S(0)$.
- Model for PS membership $S(0), S(1)$.
- Multiply likelihoods $\Rightarrow$ joint model for Y and S.
- **Treat unobserved potential outcomes as missing data** $\Rightarrow$ integrate out to define likelihood.
- Can easily include covariates in either model.
- Use (weakly informative) priors to govern "strength" of assumption, e.g. monotonicity.
- Application: Magnusson et al. (2019), Public Assessment Report of the European Medicines Agency (EPAR):
  European Medicines Agency, Committee for Medicinal Products for Human Use (2019).

Estimation approaches: principal ignorability

Principal ignorability (PI, or conditional independence):

- Approach very similar to propensity scoring in observational studies.
- Specify **separate models** for Y and S.
- Conditional on baseline covariates X: $Y(0)$ and $S(1)$ independent.
- X: all variables that **confound** $Y(0)$ and $S(1)$ $\Rightarrow$ once X are known, $S(1)$ provides no further information on $Y(0)$ (+ vice versa):

$$p(Y(0)|X, S(1)) \quad = \quad p(Y(0)|X).$$

- Allows modeling of $Y(0)$ and $S(1)$ **just based on** X. Unobserved outcome not needed in model.
- Assumption is **across worlds**.

Estimation approaches: principal ignorability

Estimand of interest:

$$P(Y(1) > t|S(1) = 1) - P(Y(0) > t|S(1) = 1).$$

Estimation:

- $P(Y(1) > t|S(1) = 1)$: survival function in ADA+ in treatment arm.
- $P(Y(0) > t|S(1) = 1)$: tricky, because $Y(0)$ and $S(1)$ **never jointly observed**.
- PI allows estimation of second quantity **just based on** X.

Randomization is key:

- Ensures that relationship $X - S$ same in both groups.
- Allows prediction of PS membership in control group using model from treatment group.

Estimation under principal ignorability for ADA example

- Estimate $P(S(1) = 1|X)$ on treatment arm using logistic regression.
- Use predicted probabilities as **weights** for patients in control arm $\Rightarrow$ make samples **comparable**.
- Compute effect measure of interest.
- Alternatives:
  - **Multiple imputation**, i.e. impute $S(1)$ for control patients. Properly accounts for uncertainty in estimated weights!
  - Plain **regression adjustment**.
  - **Matching**.
- See propensity score literature for assessment of methods, e.g. Austin (2011).

Estimation under principal ignorability for ADA example

Choice of X:

- Adjust for all confounders that make $Y(1)$ and $S(0)$ (+ vice versa) independent.
- Only adjust for X that confound Y and S across worlds: predictors of S and Y. Similar to observational studies: $X =$ predictors of treatment and outcome.
- **Do not include** covariates that "only" help predict S but have no impact on Y.
- Similar to considerations for observational studies.

Doing now what patients need next

R version and packages used to generate these slides:

R version: R version 4.0.5 (2021-03-31)

Base packages: stats / graphics / grDevices / utils / datasets / methods / base

Other packages: prodlim

This document was generated on 2021-08-24 at 09:54:02.