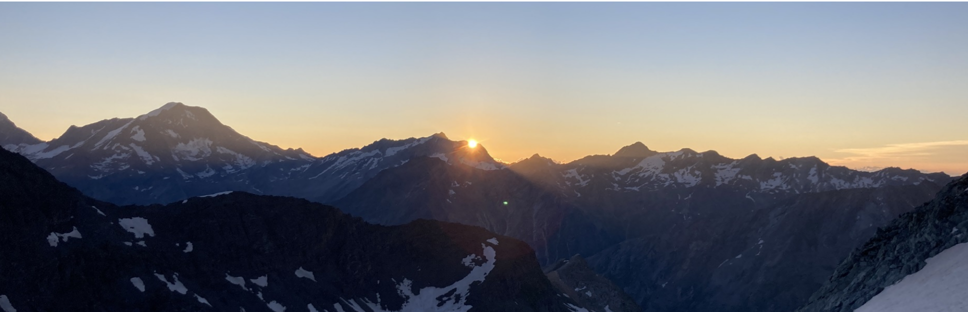
Why clinical statisticians need to know about causal inference

Kaspar Rufibach
Advanced biostatistical sciences, Merck KGaA, Darmstadt
PSI webinar 19th November 2024



Claim: clinical statisticians need to know about causal inference.

Give me the business case!

Many scientifically valid questions are currently answered making unrealistic assumptions only.

Get out of habits! Ask: Why are we doing a certain analysis?

Causal inference appears everywhere in drug development.

linkedin post

Why you need to know about causal inference.

Example 1

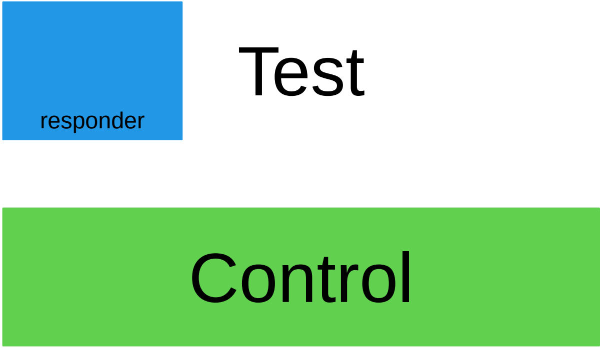
Illustration inspired by talk by Björn Bornkamp.

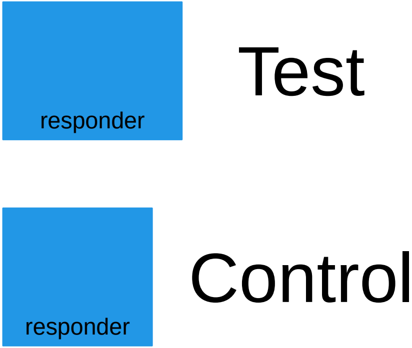
2-arm RCT test (T) vs. control (C)

Do responders have longer overall survival?

Test

Control


responder
Test
Control


responder
Test
responder
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 for cohorts built by postbaseline events are misleading and do answer causal question only under unrealistic assumptions.

Effect of response, toxicity, dose reduction, tumor growth inhibition metrics, ADA, ... on long-term endpoint.

Principal stratification estimand.

Estimation: Randomization + assumptions.

Assumptions are unverifiable.

Scientific knowledge + sensitivity analyses.

Example 2: table 2 fallacy.

A multivariate sensitivity analysis of PFS will be performed using Cox proportional-hazards regression to assess the treatment effect after adjustment for potential prognostic factors.

Table 2 fallacy: "Adjust for prognostics factors"

Outcome, **dependent**, **effect**: y.

Exposure, **independent**, **cause**: x.

Regression	**Causal inference**
"prognostic factor"	Confounder
	Mediator
	Proxy confounder
	Competing exposure

https://dagitty.net/learn/graphs/roles.html

Table 2 fallacy: "Adjust for prognostics factors"

> *...valid estimation of a causal effect requires the delineation of a range of assumptions, both causal and parametric, and there are no reasonable assumptions under which the coefficients of a multivariable regression model simultaneously provide estimates of the causal effects of every variable in the model. If this is understood , it is a short step to ask to what questions, if any, the coefficients of these ubiquitous models provide answers.*

Carlin and Moreno-Betancur (2024)

Causal inference tells us which regression model to run.

.

Example 3: How is causal inference related to the ICH estimands addendum?

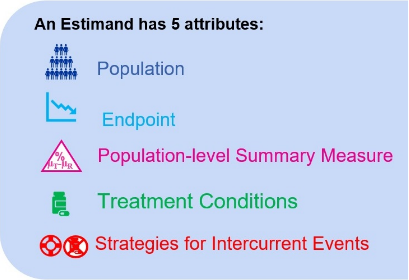
ITT?

Causality and ICH E9(R1) addendum

Defines **causal effect** (= estimand) based on potential outcomes – as does causal inference.

Addendum says nothing about **estimation** though.

Five attributes to describe effect of interest:



Past ⇒ addendum ⇒ causal inference.

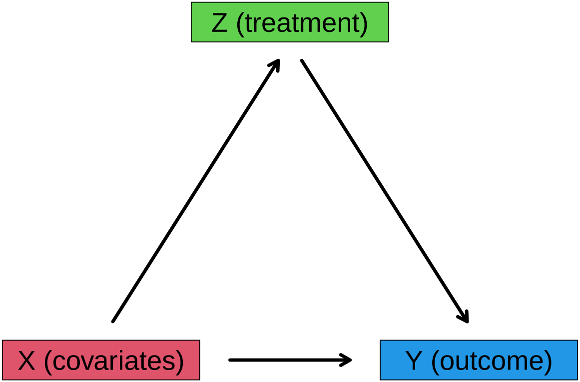
Causality and ITT

- Original ICH E9 introduced **ITT principle** (in 1998):
  - All randomized subjects included in analysis.
  - Effect of a treatment policy can be best assessed by evaluating on the basis of the **intention to treat** a subject rather than the actual treatment given.
  - Not an estimand!
  - "Principle" that refers to analysis sets $\Rightarrow$ entirely pertains to **estimation**.
- Causal inference: ITT = (population) effect of initially assigned intervention.

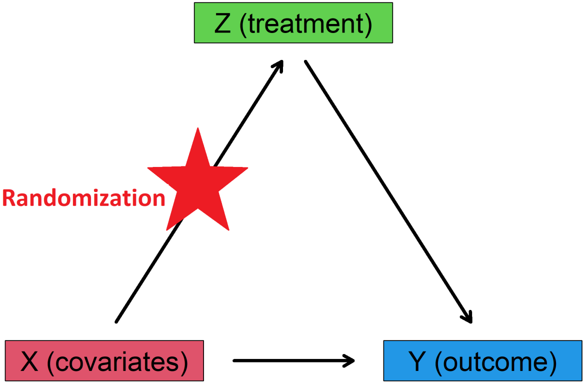
Why you need to know about causal inference.

Example 4

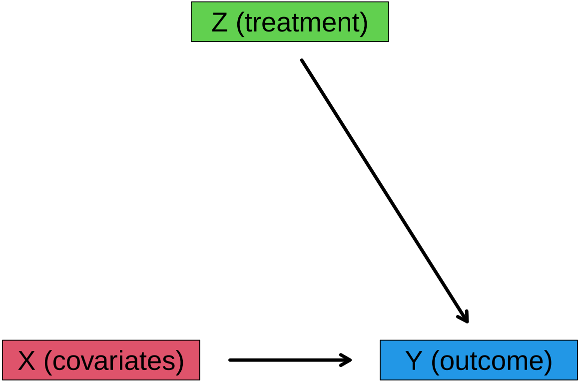
Why do we randomize?



- X: age, comorbidities, disease severity, ... **Confounder**.
- Z: experimental vs. control.
- Y: OS.



- X: age, comorbidities, disease severity, ... **Confounder**.
- Z: experimental vs. control.
- Y: OS.



Causal inference allows mathematical proof,
via potential outcomes, that randomization can be used
to obtain causal treatment effects.

Why **you** need to know about causal inference.

Example 5

Head-to-head comparison in absence of randomization using Real-world data.

Target trial emulation.

Decisions need to be made even w/o RCT – maintaining status quo is also a decision!

Target trial emulation

1. **Ask causal question**: specify protocol of analogous RCT explicitly.
2. **Answer causal question**:
    - Identify suitable RWD source(s).
    - Try and emulate RCT using RWD.

Problems with naive analysis of RWD

Asking **wrong or poorly-defined causal question**.

- Explicit representation of causal question as target trial prevents this.

Answering causal question **incorrectly**:

- Not accounting for various biases.
- Not adjusting for relevant confounders.
- Adjusting where we should not.
- Not using appropriate statistical methods.

Result: biased effect estimates, wrong conclusions.

Hernan (2020).

If RCT disagrees with RWD analysis: often, because apples compared to oranges!

Examples and concepts

Example 1: Subgroups built by post-baseline variable. **ICH E9 addendum, POs**.

Example 2: Which regression model to run. **DAGs**.

Example 3: How is causal inference related to the **ICH estimands addendum**? **ITT**?

Example 4: Why do we randomize? **Potential outcomes**.

Example 5: H2H comparison with RWD. **Target trial framework**.

Structured way to think about:

- **Generalizability**: extending causal effect from RCT to RCT's original target population.
- **Transportability**: extending causal effect from RCT to distinct population.

Many scientifically valid questions are currently answered making unrealistic assumptions only.

Get out of habits! Ask: Why are we doing a certain analysis?

Causal inference appears everywhere in drug development.

linkedin post

Thank you for your attention.

kaspar.rufibach@merckgroup.com

Slides can be downloaded on

www.kasparrufibach.ch

References I

- Carlin, J. B. and Moreno-Betancur, M. (2024). On the uses and abuses of regression models: a call for reform of statistical practice and teaching. `https://arxiv.org/abs/2309.06668`
- Hernan, M. (2020). Head-to-head comparisons using real-world data. *Presentation via Cytel*.

R version and packages used to generate these slides:

R version: R version 4.3.2 (2023-10-31 ucrt)

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

Other packages: prodlim

This document was generated on 2024-11-18 at 23:35:22.