Using principal stratification to address post-randomization events: A case study

Baldur Magnusson, Advanced Exploratory Analytics
PSI Webinar

November 2, 2017

NOVARTIS

Outline

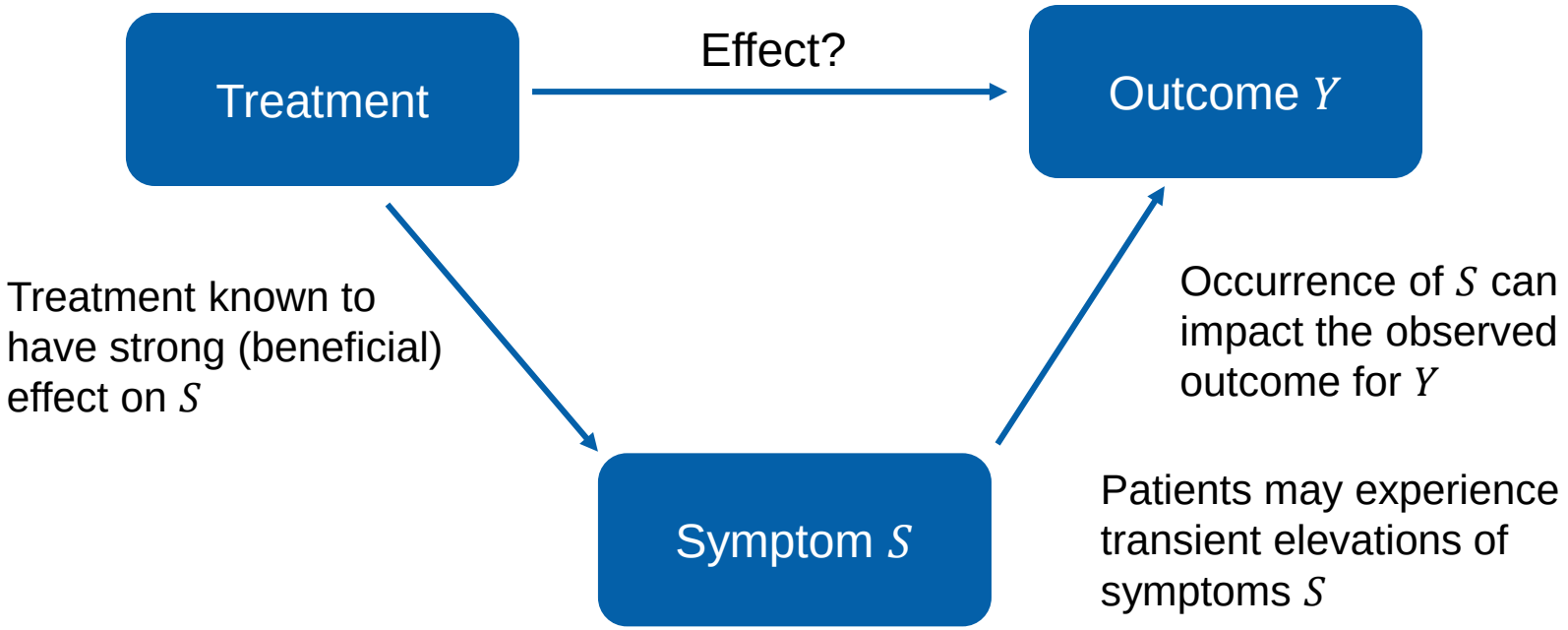
- Context
- Principal stratification
- Estimand of interest
- A glance at the Bayesian model
- Conclusions

NOVARTIS

Context

Phase 3 study with randomization to active or control

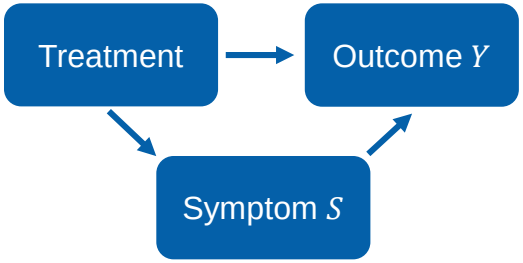
Primary question: What is the effect of treatment on outcome Y?



Question for this presentation:
How do we account for S on the estimand and the estimator level?

NOVARTIS

Context



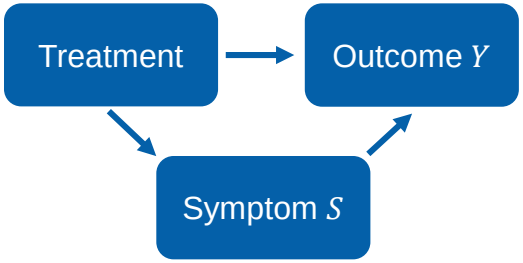
Question for this presentation:
How do we account for S on the estimand and the estimator level?

- Answer depends on the scientific question of interest...
- In our example: want to know the effect of treatment on Y among patients for whom S is very unlikely to occur
- How to reflect “very unlikely to occur” in our estimand?
- Focus on treatment effect in subgroup of patients who would not experience S regardless of treatment assignment
- This is the principal stratification estimand discussed in ICH E9 (R1)

NOVARTIS

Other estimands

As implied by common analyses



- Occurrence of S is irrelevant (treatment policy)
- Effect in population of patients without pre-study S
  - Not useful if pre-study S is not predictive of on-study S
  - Does not acknowledge the treatment effect on S
- Effect in the population of patients without on-study occurrence of S
  - Conditions on a post-randomization outcome affected by treatment
  - Estimate of treatment effect on Y would not have a causal interpretation
- Effect in a world where S would not occur
  - Hypothetical estimand since S cannot be intervened on
- None of these estimands are appropriate for our situation

NOVARTIS

Principal stratification

- Notation:
  - $S(z)$ = symptom indicator under treatment $z \in \{0,1\}$
  - That is, $S(z) = 1$ for patients who experience S if assigned to z
  - $Y(z)$ = outcome indicator under treatment $z \in \{0,1\}$
  - That is, $Y(z) = 1$ for patients who experience Y if assigned to z
- $S(z)$ and $Y(z)$ are potential outcomes
  - Every patient has a potential outcome for both $z = 0$ and $z = 1$
  - Only observe one potential outcome per patient
  - Considered as fixed attributes (baseline characteristics)
- We use S and Y to denote observed outcomes

NOVARTIS

Principal stratification

- Stratify patients as belonging to one of:
  - **Immune**: No symptom regardless of treatment
  - **Doomed**: Symptom occurs regardless of treatment
  - **Benefiter**: Symptom occurs only on placebo
  - **Harmed**: Symptom occurs only on active

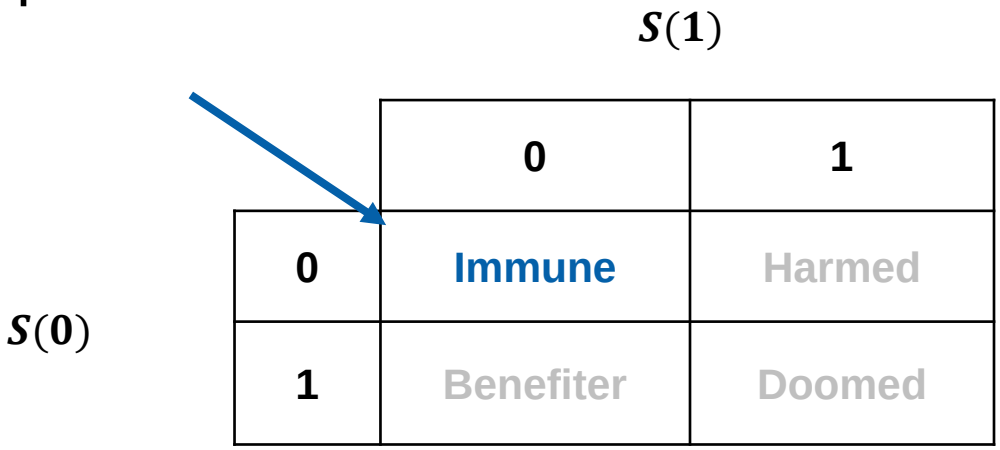
$S(0)$ \ $S(1)$	**0**	**1**
0	**Immune**	**Harmed**
1	**Benefiter**	**Doomed**

- Stratum membership not directly observable
- Observe outcome on actual treatment received
- E.g. active arm patient ($z = 1$) with $S = 0$ could be either immune or benefiter

NOVARTIS

Estimand of interest

- Interested in the difference in proportions of Y in the immune principal stratum



- Principal stratum causal effect:

$$P[Y(1) = 1 \mid S(1) = 0, S(0) = 0] \;/\; P[Y(0) = 1 \mid S(1) = 0, S(0) = 0]$$

$$= \frac{P[Y(1) = 1 \mid \text{Immune}]}{P[Y(0) = 1 \mid \text{Immune}]}$$

NOVARTIS

Identifying the estimand

Assumptions

- In practice, only observe the margins from this table
- Need identifying assumptions in order to link estimand to 'observables'

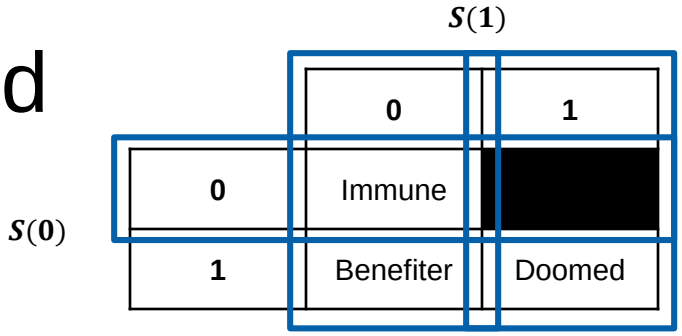
$S(0)$ \ $S(1)$	0	1	Sum
0	??		✓
1	??	??	✓
Sum	✓	✓	

- Monotonicity assumption: There are no harmed patients
  - A patient not experiencing S on placebo will not experience S on active
  - That is, $S(0) = 0 \Rightarrow S(1) = 0$
  - A patient experiencing S on active will experience S on placebo
  - That is, $S(1) = 1 \Rightarrow S(0) = 1$

NOVARTIS

Identifying the estimand

Principal strata proportions



- Monotonicity allows some patients to be classified
  - Placebo patient with $S(0) = 0$ must be immune
  - Treated patient with $S(1) = 1$ must be doomed
- Some patients remain not classifiable
  - A treated patient who does not experience a symptom, i.e. $S(1) = 0$, could be immune or a benefiter
- We can now estimate the strata proportions
  - $P[\text{Doomed}] = P[S = 1 \mid Z = 1]$
  - $P[\text{Immune}] = P[S = 0 \mid Z = 0]$
  - $P[\text{Benefiter}] = 1 - P[\text{Doomed}] - P[\text{Immune}]$

NOVARTIS

Identifying the estimand

- Estimand of interest:

$$\frac{P[Y(1) = 1 \mid \text{Immune}]}{P[Y(0) = 1 \mid \text{Immune}]}$$ ✘ ✓

- Randomization and monotonicity allow us to identify the denominator:

$$\mathrm{P}[Y(0) = 1 \mid \text{Immune}] = P[Y = 1 \mid Z = 0, S = 0]$$

- Because $S(1) = 0$ could imply immune *or* benefiter, the numerator is not identifiable
- However, bounds on the numerator can be derived leading to a range of feasible values for the estimand

NOVARTIS

Identifying the estimand

- Using the law of total probability and without further assumptions:

$$P(Y(1)=1|I) = \underbrace{\frac{P(Y(1)=1|I \text{ or } B)}{P(I|I \text{ or } B)}}_{\text{Intercept}} - \underbrace{\frac{P(B|I \text{ or } B)}{P(I|I \text{ or } B)}}_{\text{Slope}} P(Y(1)=1|B)$$

- Intercept and slope can be calculated from the data
  - $P[Y(1)=1 \mid \text{Immune or Benefiter}] = P[Y=1 \mid Z=1, S=0]$
  - $P[\text{I} \mid \text{I or B}]$ is a function of strata proportions
- $P[Y(1)=1 \mid \text{Benefiter}]$ cannot be identified
  - Known to be between 0 and 1
  - Could make further assumptions, e.g.

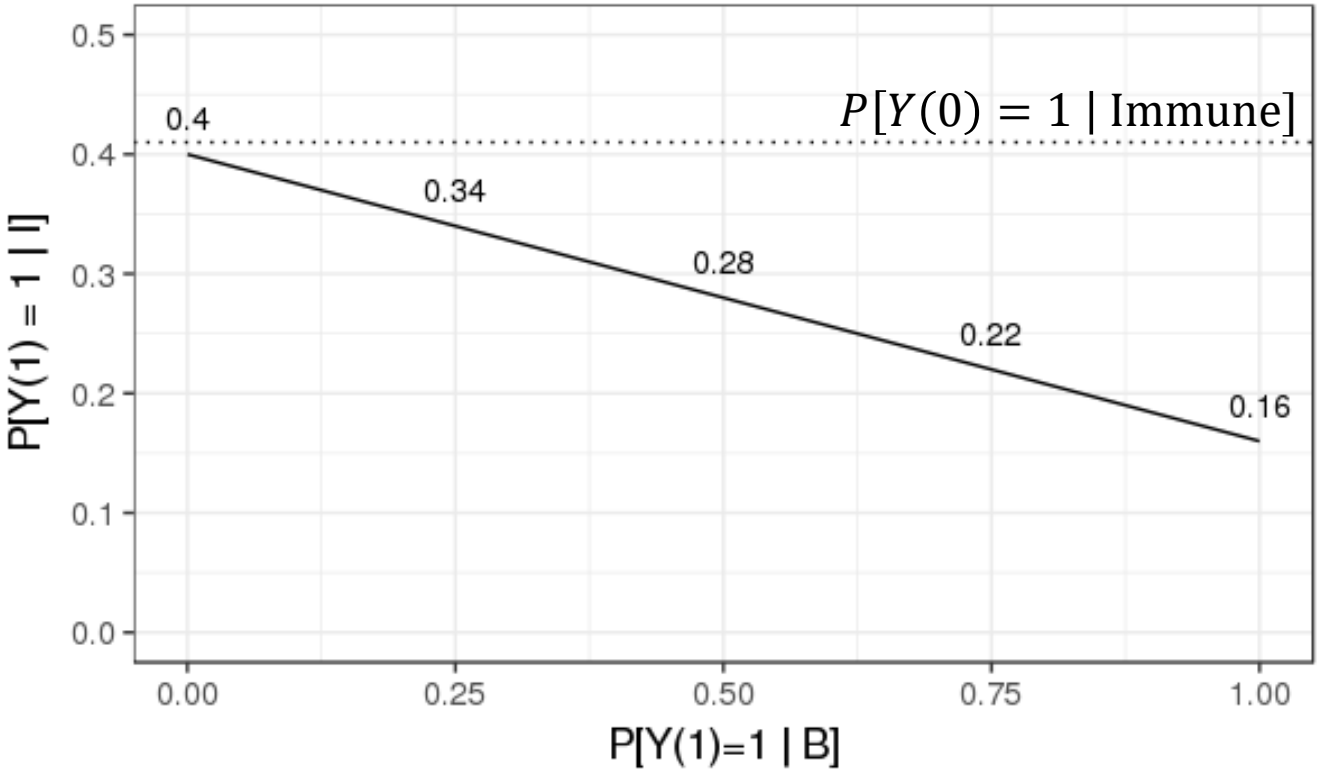
$$P[Y(1)=1 \mid \text{Benefiter}] \leq P[Y(1)=1 \mid \text{Doomed}]$$

NOVARTIS

Identifying the estimand

Visualizing a range of feasible values

- We can calculate $P[Y(1) = 1 \mid \text{Immune}]$ for a range of values of $P[Y(1) = 1 \mid \text{Benefiter}]$



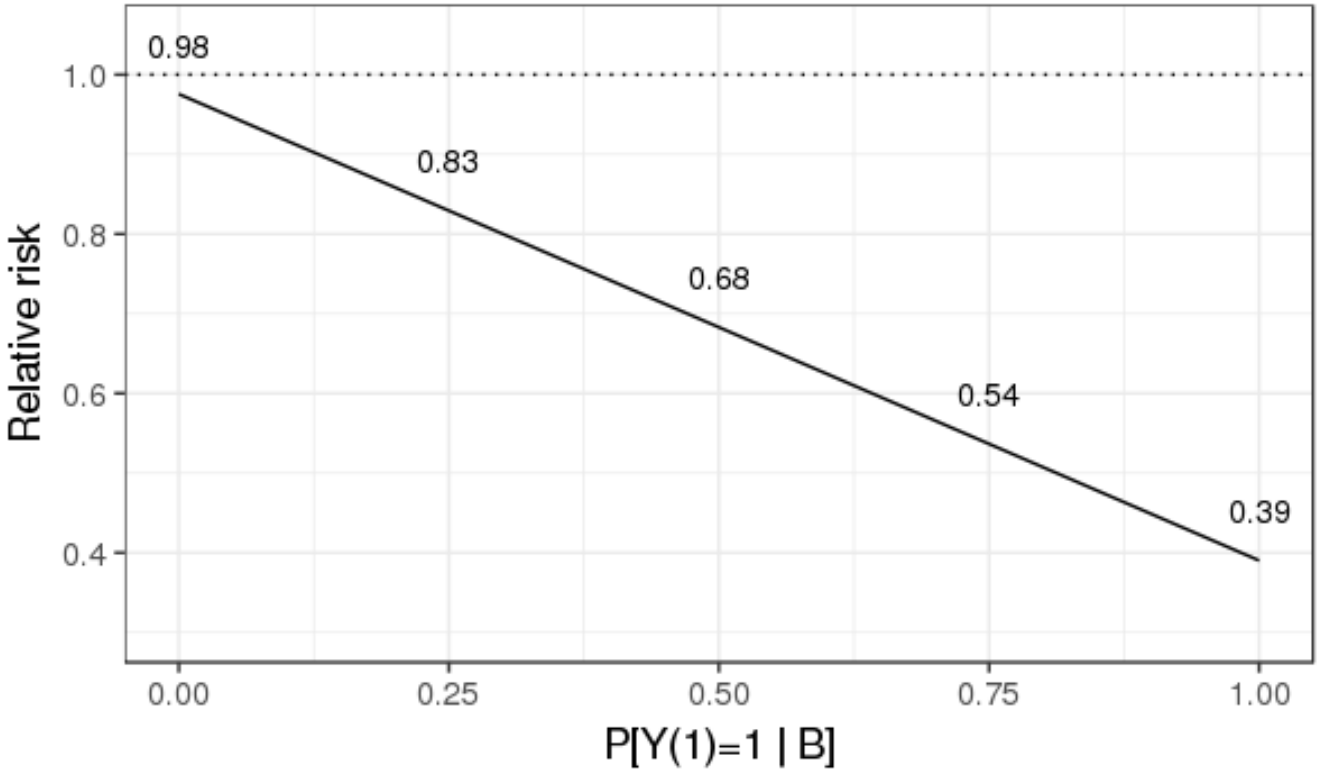
*simulated data

NOVARTIS

Identifying the estimand

Visualizing a range of feasible values

- We can also calculate the range of feasible values for the estimand of interest



*simulated data

NOVARTIS

Estimation

- So far... no estimation, no uncertainty
- It is straightforward to estimate the parameters in the Bayesian framework (e.g. using Stan)
- Could use the equation for $P[Y(1) = 1 \mid \text{Immune}]$:

$$P(Y(1)=1|I) = \underbrace{\frac{P(Y(1)=1|I \text{ or } B)}{P(I|I \text{ or } B)}}_{\text{Intercept}} - \underbrace{\frac{P(B|I \text{ or } B)}{P(I|I \text{ or } B)}}_{\text{Slope}} P(Y(1)=1|B)$$

- May result in negative values, so preferable to code the likelihood directly

Estimation

Simplified glance at the Bayesian model

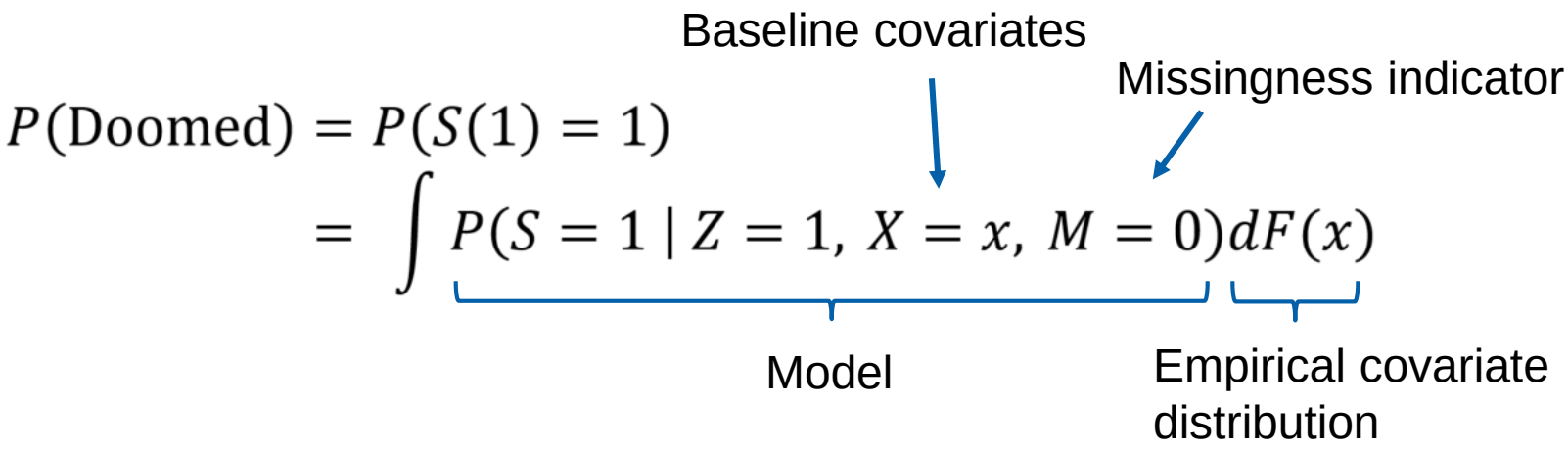
- Principal strata proportions:
  - $S \mid Z = 0 \sim \text{Bernoulli}(1 - \pi_{Immune})$
  - $S \mid Z = 1 \sim \text{Bernoulli}(\pi_{Doomed})$
  - $\pi_{Benefiter} = 1 - \pi_{Immune} - \pi_{Doomed}$
- Outcome model:
  - $Y \mid S = 0, Z = 0 \sim \text{Bernoulli}(\theta_{Immune,\, placebo})$
  - $Y \mid S = 0, Z = 1 \sim \frac{\pi_{Immune}}{\pi_{Immune} + \pi_{Benefiter}} \text{Bernoulli}(\theta_{Immune,\, active}) + \frac{\pi_{Benefiter}}{\pi_{Immune} + \pi_{Benefiter}} \text{Bernoulli}(\theta_{Benefiter,\, active})$
- Results summarized by examining the posterior distribution of $\theta_{Immune,\, active} / \theta_{Immune,\, placebo}$
- Sensitivity analyses:
  - Partially relax monotonicity assumption (e.g. through a strongly informative prior)
  - Explore various informative priors for $\theta_{Benefiter,\, active}$

NOVARTIS

Estimation

Covariates and missing data

- Due to variable follow-up time, not all patients have available S and Y data in the time period of interest
- Address this using standardization
- Write standardized probability of interest as



Conclusions

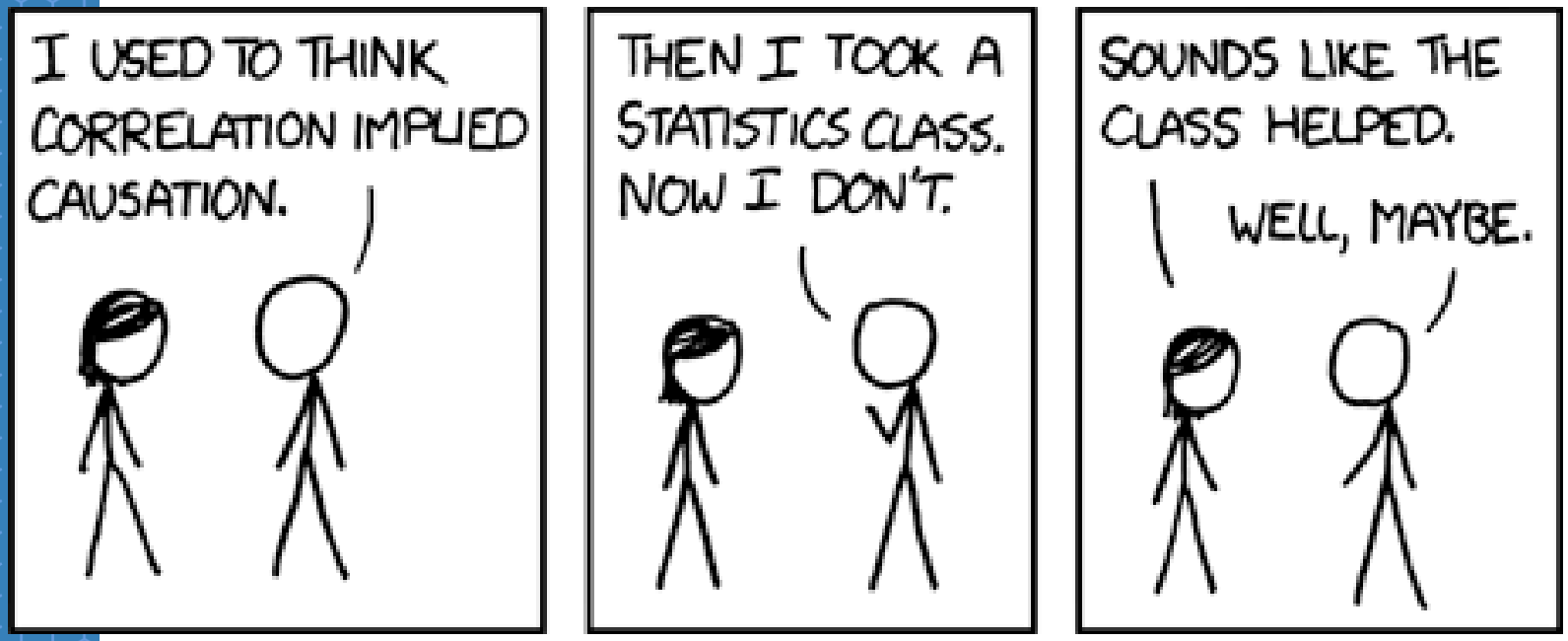
- Causal inference framework (potential outcomes) provides a natural way of defining the estimand of interest
- Principal stratification is not the *only* way to approach this type of problem...
  - The appropriate estimand depends on the specific scientific objectives
- Monotonicity assumption is critical
  - Substantive rather than statistical – ideally backed by strong clinical rationale
- Bayesian framework is appealing in this setting:
  - Straightforward to model principal strata proportions
  - Use of mixture distribution to handle lack of identifiability in the active arm
  - Can explicitly encode our (lack of) knowledge about certain parameters using prior distributions

NOVARTIS

Acknowledgments

- Dan Scharfstein
- Heinz Schmidli
- Frank Bretz
- Sebastian Weber
- Nicolas Rouyrre
- Nikos Sfikas
- David Ohlssen

NOVARTIS



https://xkcd.com/552/

Thank you