Implementing the Principal Stratum estimand strategy using Instrumental Variable methods: An emulation of the CANTOS trial

Jack Bowden

August 2019

MRC Integrative Epidemiology Unit

Causal inference in trials, a popular view

- *'The language, terminology or notation of causal inference is unnecessary in RCTS'*
- *'Randomization allows us to unambiguously interpret treatment effect estimates as causal estimates (if you must use that word)'*
- So do we really need DAGs, potential outcomes & fancy causal inference methods?
- As a statistician working in Epidemiology for the last 6 years, I want to convince you that these tools are useful for trials and the **Estimand Framework**
- Focus on Instrumental Variable methods

Before I get to the CANTOS trial...

- I will try to give an explanation of

  - ITT analysis
  - Principal Stratification
  - IV regression

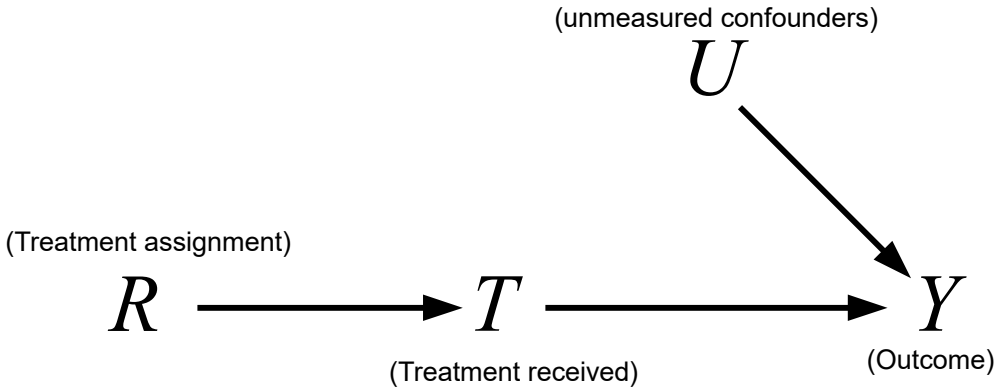
- Explain the danger of being seduced by Principal Stratification

- Try to clarify the interpretation of causal estimates in the presence of **treatment effect heterogeneity**

  - & how this assumption can be relaxed with extended two-parameter causal models

- So please bear with me, it has real relevance to CANTOS!

An idealised RCT



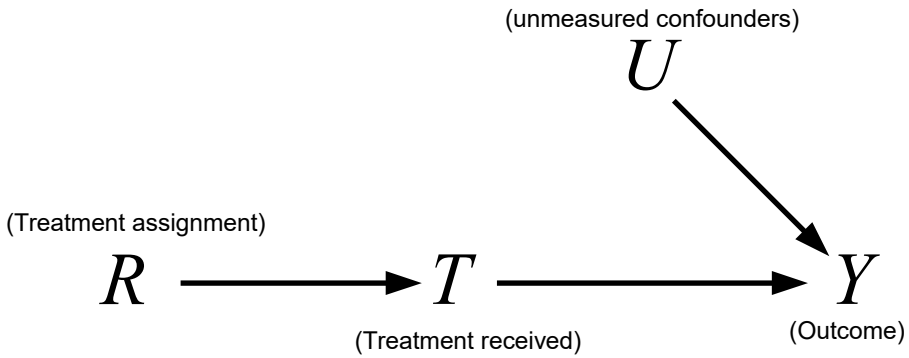
- Random assignment perfectly predicts treatment received, so nothing else can
- A standard comparison of patient outcomes across randomized groups tells us about the **average causal effect** of treatment

$$\text{ITT effect} = E[Y|R=1] - E[Y|R=0] = \text{causal effect}$$

- Using the potential outcomes notation

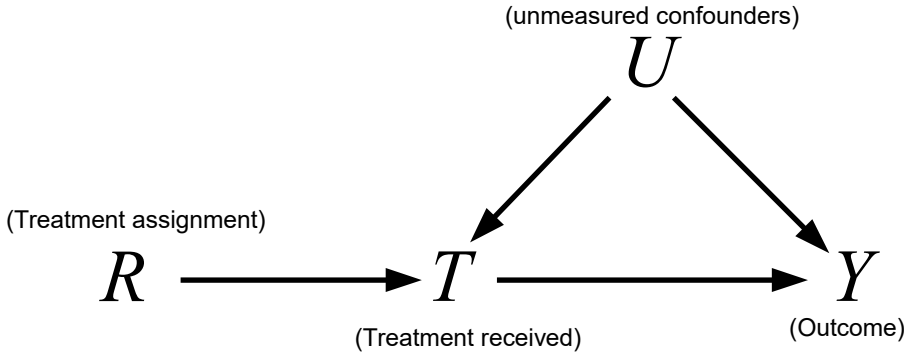
$$E[Y(R=1) - Y(R=0)]$$

Randomization is the ultimate Instrumental Variable



- IV1: Randomization predicts treatment
- IV2: Randomization is independent of everything else
- IV3: Randomization only affects patient outcomes via treatment

Randomization is the ultimate Instrumental Variable



- IV1: Randomization predicts treatment
- IV2: Randomization is independent of everything else
- IV3: Randomization only affects patient outcomes via treatment
- What if (potentially unmeasured) post-randomization factors could influence patient adherence to the treatment & the outcome?
- Standard as treated and per-protocol analyses will be biased
- Adjust for all confounders in the ITT analysis? Yes, if possible
- But, randomization is **still** a valid IV

Causal inference in trials: the Estimand Framework

- The ICH E-9 Addendum is forcing trialists to be much more forward thinking and upfront about the issue of **Intercurrent Events**
- An intercurrent event is

  '*any event occurring between the initial randomization of a patient and the observation of their final outcome which complicates the description and interpretation of the treatment effect*'

- Trialists must have an 'Estimand Strategy'
- So what strategies do IV methods have a role in delivering?
- *Estimands and Sensitivity Analysis in Clinical Trials*. ICH Harmonised guideline 2017

Causal inference in trials: the Estimand Framework

- **Treatment policy strategy**: *Occurrence of the intercurrent event is deemed to be irrelevant, all patient outcomes are used regardless of whether the intercurrent event occurred or not*

  - Can be obtained via an ITT analysis

- **Principal Stratum strategy:** *Estimate the treatment effect in a target population for whom the intercurrent event would not occur*

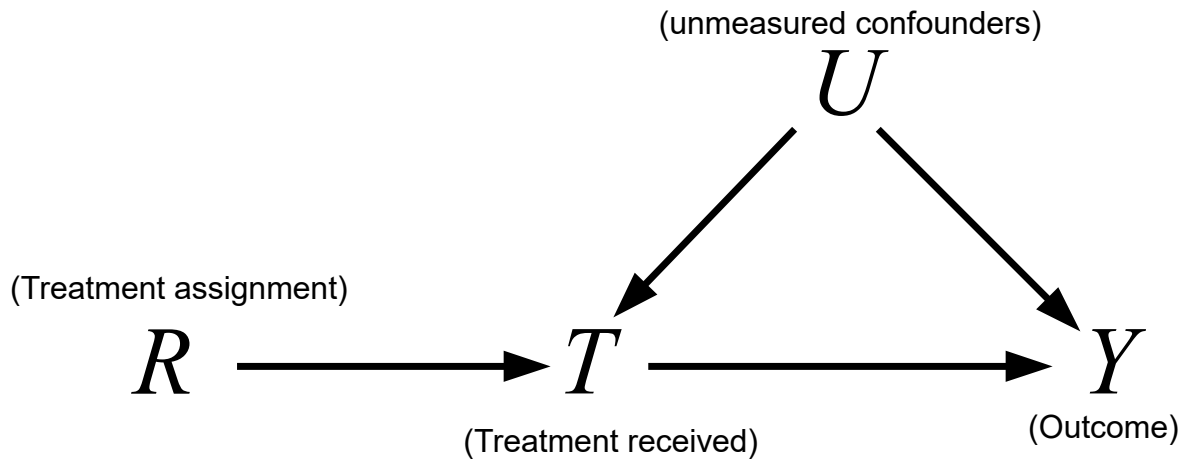
  - Most naturally obtained using Principal Stratification

- Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics*. 2002;58:21–9.

- **Hypothetical strategy**: *Estimate the outcome variable (and from that the treatment effect) under the hypothetical scenario in which the intercurrent event did not not occur*

  - IV methods?

Example: Intercurrent Event = classic non-adherence



- Assume R and T are binary (0,1) variables
- Treatment non-compliance means R may not equal T
- Common set up in academic world
- Assume outcome continuous/binary & trt effect = mean/risk diff.

Compliance classes using potential outcomes

- Define the four Principal Strata
- **Never Takers**: $T(R = 0) = T(R = 1) = 0$
- **Always Takers**: $T(0) = T(1) = 1$
- **Compliers**: $T(0) = 0, T(1) = 1.$
- **Defiers**: $T(0) = 1, T(1) = 0$
- The Principal Stratum estimand might then be the ITT effect within **compliers**
  - $E[Y|R = 1, T(0) = 0, T(1) = 1] - E[Y|R = 0, T(0) = 0, T(1) = 1]$
  - Referred to as the Complier Average Causal Effect (CACE)

Identification of the CACE

Assume no defiers (or monotonicity)

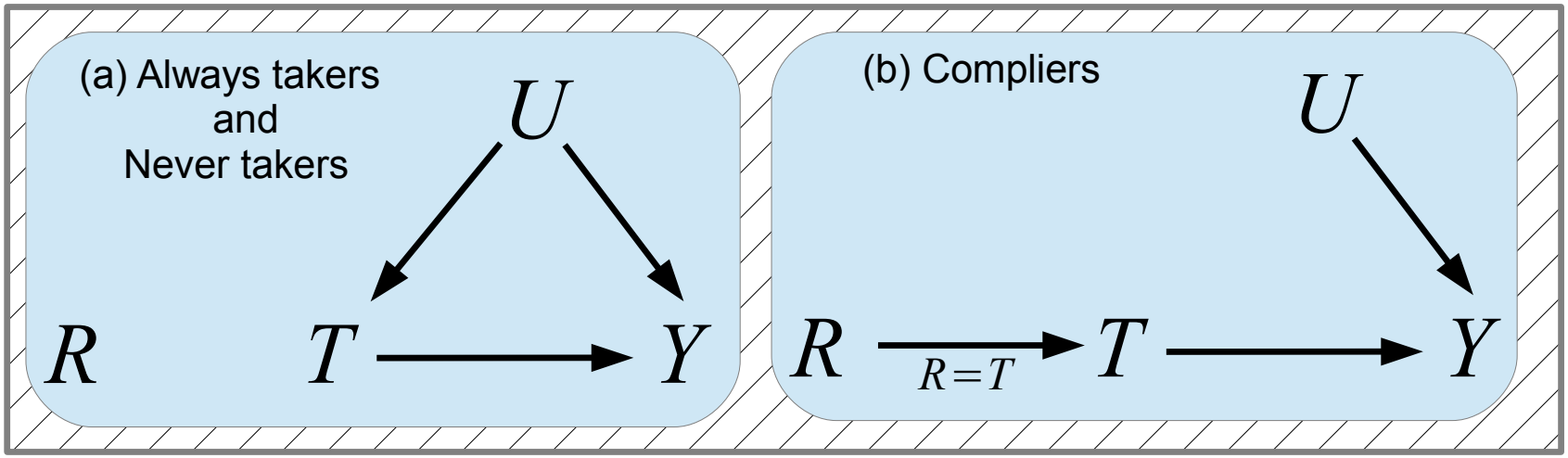
	$T(R=1)$	$T(R=0)$	Proportion	Estimated by	Who?
Always Takers	1	1	π_{at}	$Pr(T=1\|R=0)$	Control arm non-adherers
Never Takers	0	0	π_{nt}	$Pr(T=0\|R=1)$	Treatment arm non-adherers
Compliers	1	0	π_c	$1-\pi_{at}-\pi_{nt}$	*Complier fraction*
Defiers	0	1	π_d	Fixed at zero	

From this we can identify the complier fraction π_c

The CACE can be shown to equal $\dfrac{E[Y|R=1]-E[Y|R=0]}{\pi_c}$

$$= \frac{\text{ITT effect}}{\text{Complier fraction}}$$

DAG intuition



- Randomization is a perfect IV in Compliers

  - ITT effect = causal effect of treatment

- Randomization fails the most basic IV test (of not predicting treatment) in the Always and Never Takers
- Doesn't mean Always and Never Takers would experience a zero treatment under hypothetical intervention
- It does mean their ITT effect is zero

An alternative Hypothetical estimand

- *'What would be the effect, among the treated population, of intervening and setting their treatment level to zero?'*
- Called the 'average effect of treatment in the treated' (ATT)
- We could write this estimand as

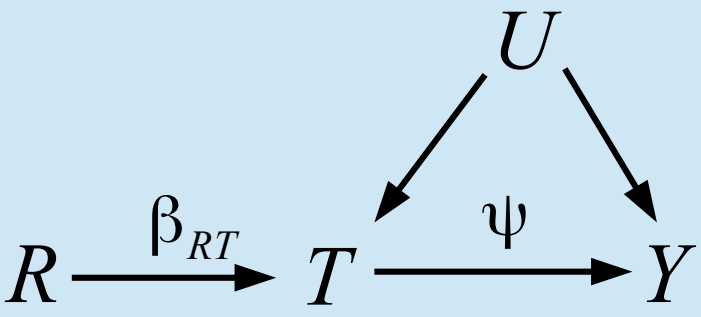
$$E[Y - Y(T = 0)|T = t] = \psi t$$

- In Principal Stratification parlance, ψ represents the causal effect across **Compliers & Always Takers** (under monotonicity)
- It assumes treatment effect homogeneity
- We can estimate ψ via G-estimation, but also using **IV regression**

Estimation using IV regression

IV regression

IV estimation: Ratio method



$$\beta_{RY} = \psi \beta_{RT}$$

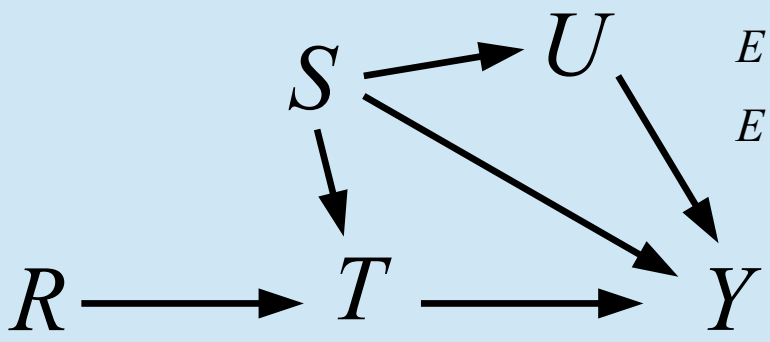
$$\psi = \frac{\beta_{RY}}{\beta_{RT}} = \frac{Cov(R,Y)/Var(R)}{Cov(R,T)/Var(R)}$$

$$\psi = CACE$$

- Perform linear regression of T on R to obtain $\hat{\beta}_{RT}$
- Perform linear regression of Y on R to obtain $\hat{\beta}_{RY}$
- Calculate the ratio $\frac{\hat{\beta}_{RT}}{\hat{\beta}_{RY}}$
- The CACE estimate equals the ATT estimate
- $\hat{\beta}_{RT}$ equals the Complier fraction estimate $\hat{\pi}_c$

Generalization to Two-Stage Least Squares

IV regression: TSLS incorporating a covariate S



$$E[Y|T,S,U]=\beta_0+\psi T+\beta_s S+U$$ **Outcome model**

$$E[T|R,S]=\hat{T}=\beta_{R0}+\beta_{RT}R+\beta_{ST}S$$ **1st stage model**

$$E[Y|\hat{T},S]=\beta_0+\psi\hat{T}+\beta_{SY}S$$ **2nd stage model**

$\hat{\beta}$ **from 2nd stage = causal estimate**

- Can seamlessly incorporate covariates and multi-valued treatments
- **step 1:** Linear regression of T on R and S to give $\hat{T}$
- **step 2:** Linear regression of Y on $\hat{T}$ and S

  - Take the coefficient of $\hat{T}$ from the model as estimate for ψ

Monotonicity and Homogeneity

- Monotonicity enables us to identify the CACE, but not the Compliers
- So is Monotonicity a useful assumption?
- The CACE/IV estimate is equal to average effect of treatment in the treated under treatment effect homogeneity
- The treated population are easily identifiable and are a larger group

  - If you were a pharmaceutical company or regulator, which group would you rather approve the a drug for?
- So is treatment effect homogeneity a useful assumption?

Monotonicity and Homogeneity

- Monotonicity enables us to identify the CACE, but not the Compliers
- So is Monotonicity a useful assumption?
- The CACE/IV estimate is equal to average effect of treatment in the treated under treatment effect homogeneity
- The treated population are easily identifiable and are a larger group

  - If you were a pharmaceutical company or regulator, which group would you rather approve the a drug for?
- So is treatment effect homogeneity a useful assumption?
- Is it reasonable to assume treatment effect homogeneity?
- What happens if it doesn't hold, and what can we do about it?

Relaxing the treatment effect homogeneity assumption

- We can estimate a separate treatment effect for Compliers and Always-Takers
- Re-write the estimand as

$$E[Y - Y(T = 0)|T = t, R = r] = \psi tr + \psi^* t(1 - r)$$

- ψ now represents the treatment effect in Compliers+Always Takers
- ψ^* represents the treatment effect in Always Takers
- Both parameters are identifiable if baseline covariates exist which

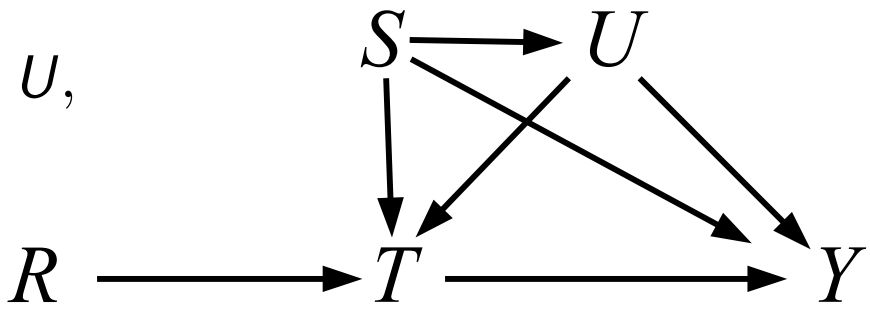
  - 1: Differentially predict treatment across arms

  - 2: Do not directly modulate the treatment effect

- Exploit by incorporating $R \times S$ interaction in 1st stage TSLS model

Proof of concept simulation (Continuous outcome)

$$
\begin{aligned}
R, S &\sim \text{Bern}(0.5) \\
U &\sim N(0, 0.5) + 0.1S \\
\eta_T &= -2 + 2R + 2RS + U, \\
P_T &= \frac{\exp(\eta_T)}{1 + \exp(\eta_T)} \\
T &\sim \text{Bern}(P_T) \\
Y &= 100 - 3TR - 2T(1 - R) + U + S + \epsilon_y, \epsilon_y \sim N(0, 1)
\end{aligned}
$$

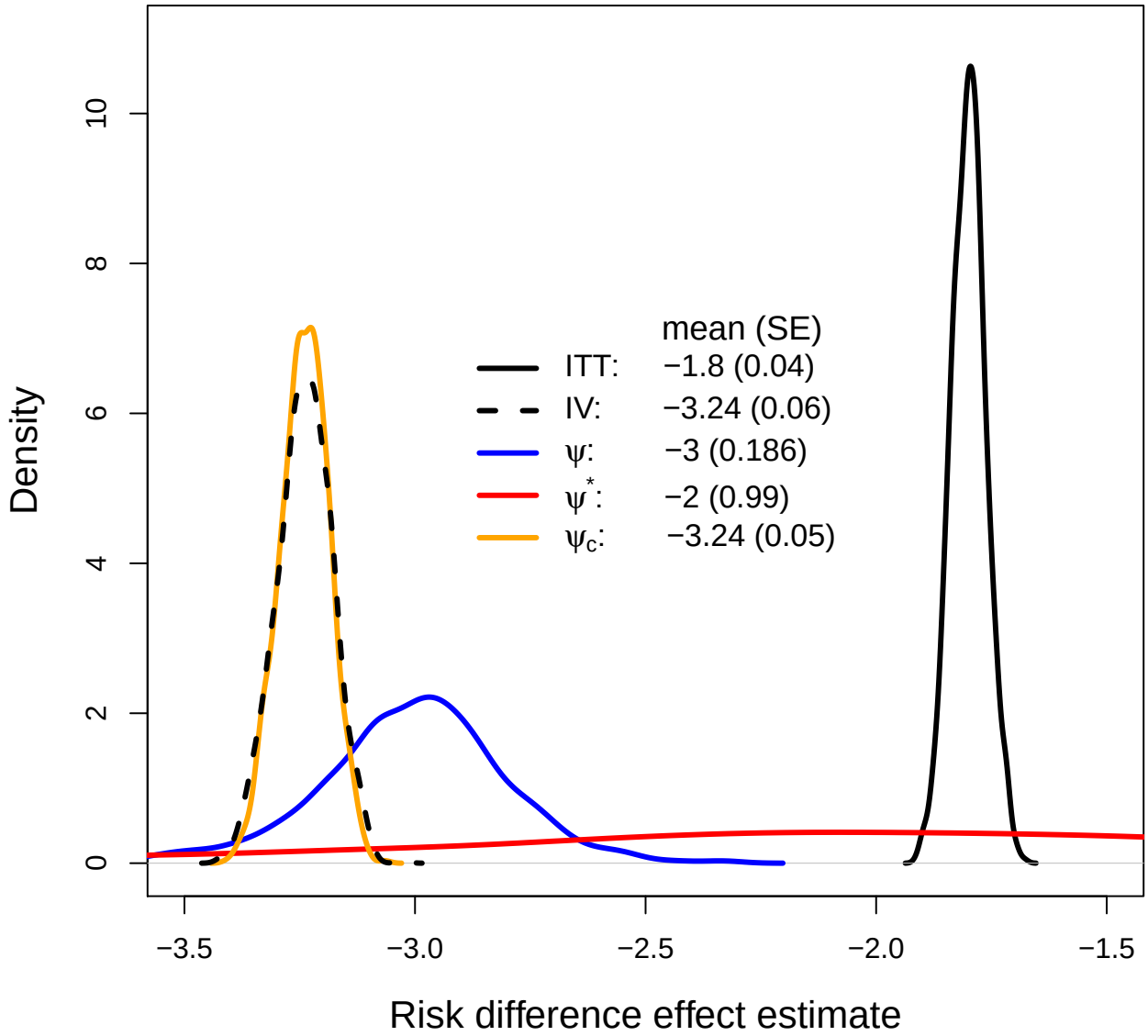


- We allow for treatment effect heterogeneity by setting

  - ψ=-3
  - ψ^*=-2.

- This implies a treatment effect in Compliers of ψ_c = -3.24

- ψ_c is a function of ψ and ψ^*



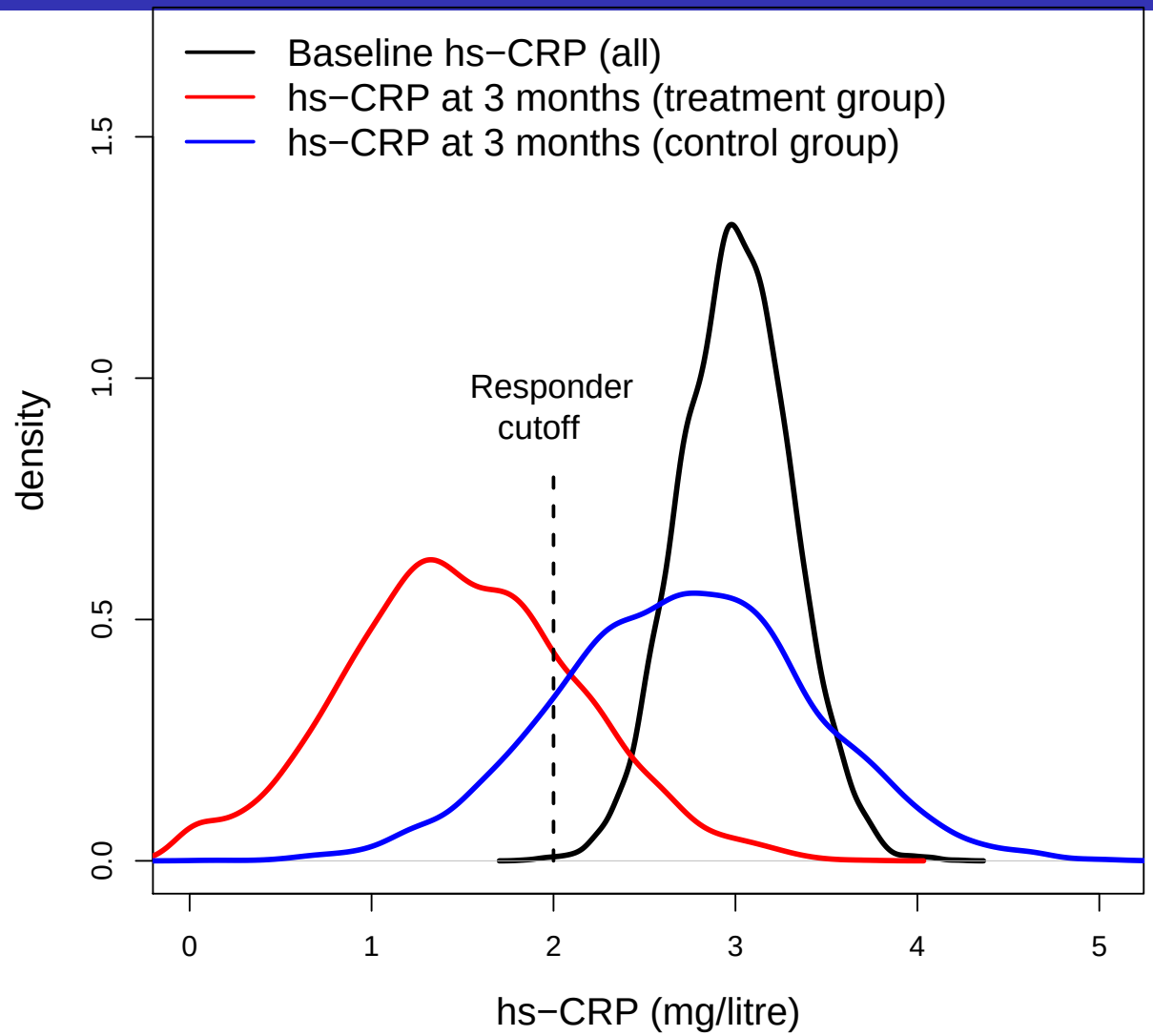
- Basic IV regression still identifies the treatment effect in Compliers
- Estimates for Treated and Always Taker pop^{ns} relatively imprecise

Application to the CANTOS trial Ridker (2017), Bornkamp and Bremen (2019)

- The CANTOS trial sort to test whether Canakinumab, an antibody which acts to reduce inflammation, was effective in reducing the risk of a major cardiac event in over 10,000 patients.
- **Eligibility**: patients must have had a previous myocardial infarction and evidence of inflammation

  - as measured by a *hs*-CRP $\geq$ 2mg/litre in their blood
- After 48 months treatment group experienced a 60% reduction in *hs*-CRP levels (control group 20%)
- Overall survival in the treatment groups was higher than the placebo group

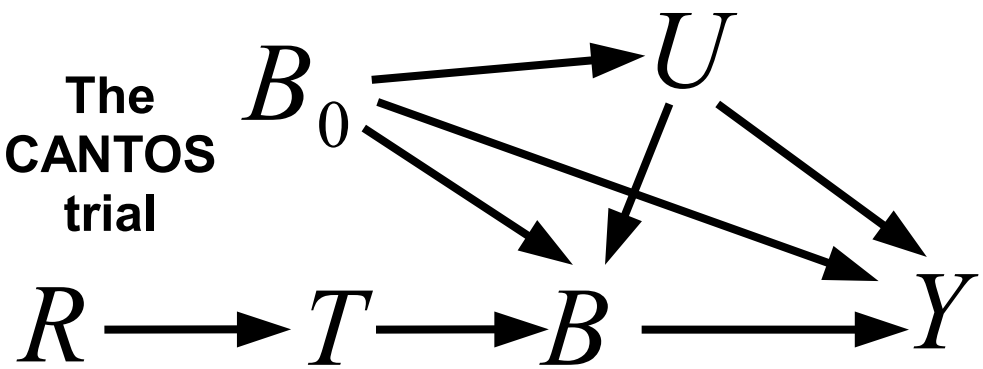
An emulation of the *hs*-CRP data after 3 months



- Some people 'respond' to treatment wrt CRP, some don't
- 'Respond' = *hs-CRP* ≤ 2 after 3 months

A Principal Stratum estimand for the CANTOS trial

- Suppose we believe the treatment operates through biomarker *hs*-CRP
- No response in *hs*-CRP $\Rightarrow$ no effect of treatment
- B = 'Biomarker responder' status (1 if *hs*-CRP ≤ 2, 0 otherwise)
- **The intercurrent event**: Biomarker non-response
- Define B_0 as baseline *hs*-crp



- Note $R = T$

A Principal Stratum estimand (Bornkamp and Bremen, 2019)

- Interested in the ITT effect in 'Biomarker responders'

  - Patients who, if they had been randomized to receive the treatment, would be biomarker responders

- Define $B(R = 0) = B(0)$ and $B(R = 1) = B(1)$ as potential biomarker response variables

- **Their estimand**:

  $E[Y|R = 1, B(1) = 1] - E[Y|R = 0, B(1) = 1]$

- Estimated $E[Y|R = 1, B(1) = 1]$ from treatment arm biomarker responders
- Imputed $E[Y|R = 0, B(1) = 1]$ by adjusting for all* confounders of response and outcome

- Is it possible to use IV methods instead?

Obtaining the Principal Stratum estimand using IV approaches

- The population of interest is the union of

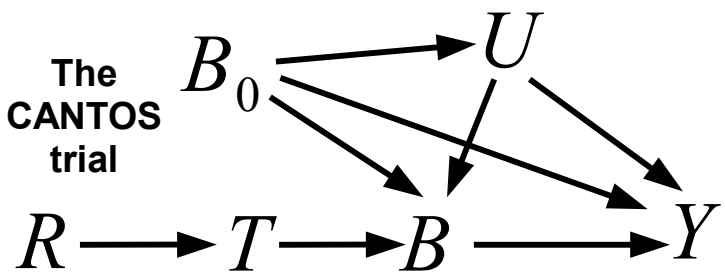
  - Always Responders: Those with $B(0) = 1, B(1) = 1$
  - 'Compliers': Those with $(B(0) = 0, B(1) = 1)$

- Analagous to the 'treated' population in previous example
- Define treatment effect in Always responders + Compliers as ψ
- Define treatment effect in Always responders as ψ^*
- The Principal Stratum estimand equals

$$\psi Pr(B = 0|T = 0) + (\psi - \psi^*)Pr(B = 1|T = 0)$$

- Under treatment effect homogeneity, 2nd term disappears
  - Standard IV estimate valid
- Use Baseline *hs*-CRP as interacting covariate to avoid this assumption

CANTOS trial data emulation

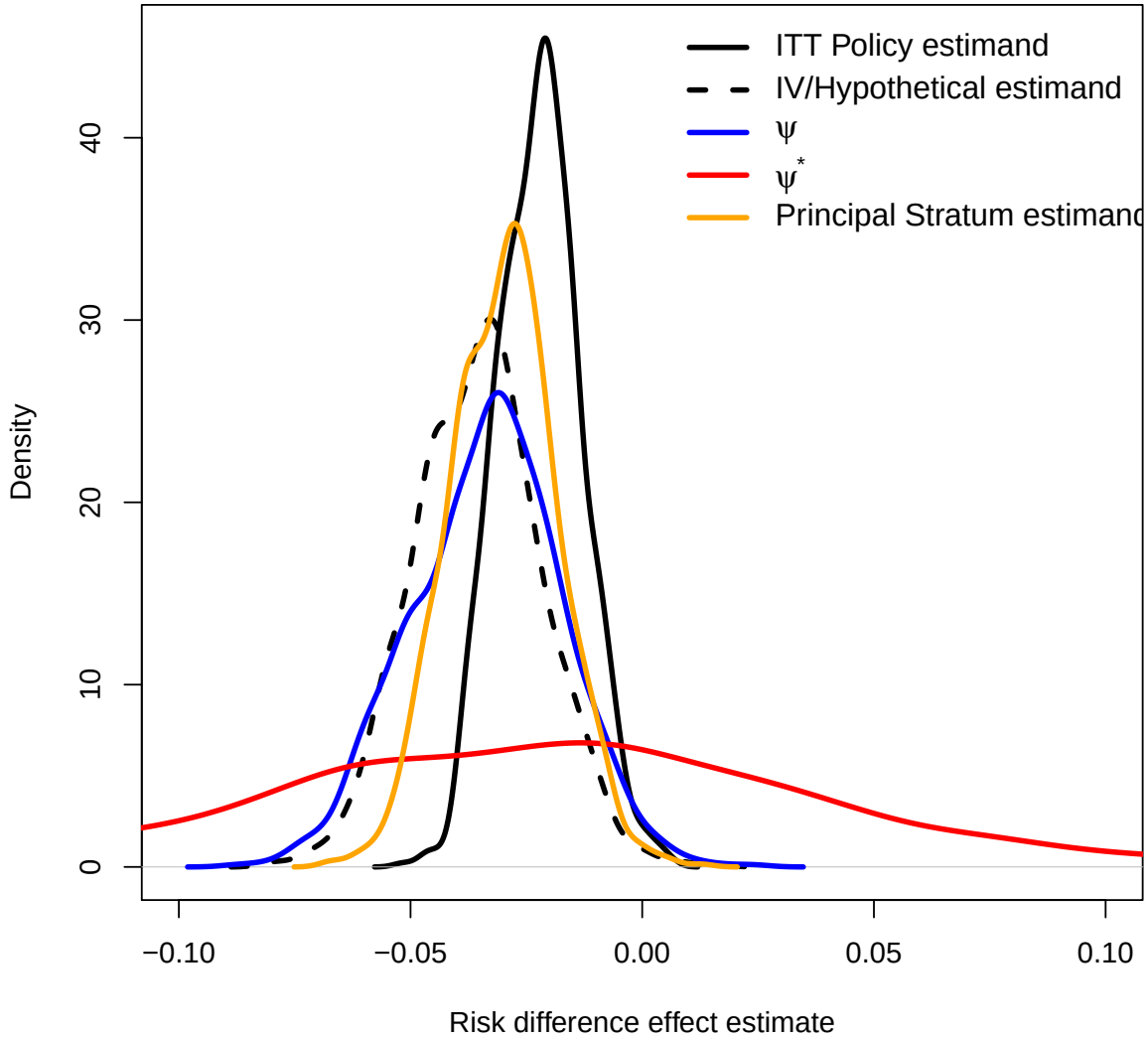


- Trial data simulated from

$$Pr(Y = 1) = 0.155 + \psi BR + \psi^* B(1 - R) + \alpha_1 U + \alpha_2 B_0 + \epsilon_Y$$

- ψ set to -0.035 and ψ^* set to -0.025
- 3000 patients per-arm
- Mean mortality rate = 16% in controls, 14% in the treatment arm
- 77% of treatment arm and 16% of control arm were biomarker responders

Results



- Principal Stratum estimand remarkably precise
- Can be explained by strong correlation between $\hat{\psi}$ and $\hat{\psi}^*$

Summary 1

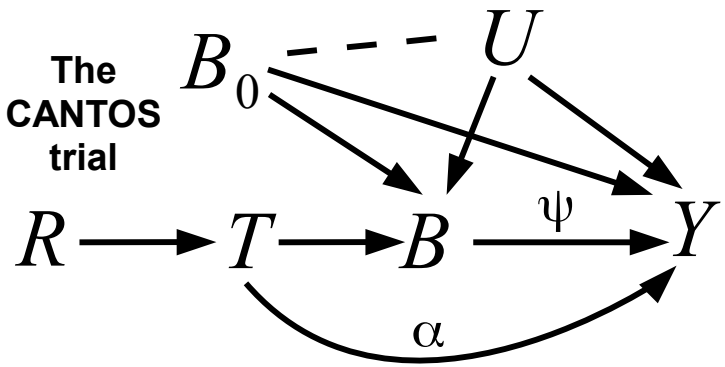
- Many different Hypothetical and Principal Stratum estimands yield the **same estimate** unless treatment effect heterogeneity explicitly modelled
- However, the resulting two-parameter causal models are more challenging to fit
- Principal Stratification is arguably responsible for simplistic dichotomization of treatment

  - I even did it myself!
- Would like to revisit the CANTOS trial and treat Biomarker response as a continuous variable

  - Not possible within Principal Stratification, but easy for IV regression
- But how to phrase within current Estimand Framework guidance?

Summary continued: Other uses for baseline covariates

- Instead of using Baseline *hs*-CRP to relax the treatment effect heterogeneity assumption we could *assume homogeneity* and estimate the direct and indirect effect of treatment



- Principal Stratum Estimand $= \alpha + \psi Pr(B = 0 | T = 0)$ in this context
- Mechanism of action hypothesis true $\Rightarrow \alpha = 0$
- Equivalent to testing and adjusting for violation of the Exclusion restriction
- A popular analysis in Epidemiological circles
- Simulations show that the Principal Stratum estimand can be estimated precisely, because $\hat{\alpha}$ and $\hat{\psi}$ are negatively correlated

Some references

- Ridker et al, (2017). Anti-inflammatory Therapy with Canakinumab for Atherosclerotic Disease *NEJM* 2017; **377**:1119–1131

- Bornkamp B, Bermann G. 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* 2019

- Bowden J, Bornkamp B, Glimm E, Bretz F. Estimating treatment effects using instrumental variable methods: from 'academic' causal inference to the Estimand Framework. *Manuscript in preparation* 2019

G-estimation of ψ

Structural mean model for the treated:

(1) Unidentified SMM: $E[Y - Y(0)|T, R] = \psi TR + \psi^* T(1-R)$

(2) Identified SMM: $E[Y - Y(0)|T, R] = \psi T$

$\psi = \psi^*$

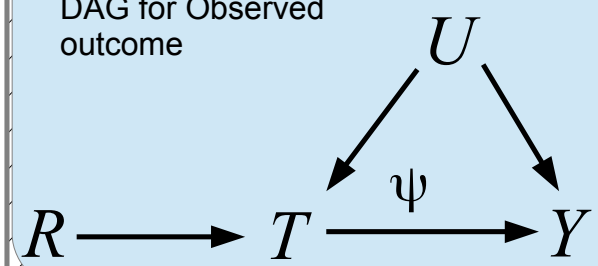
(3) **G-estimation:**

$$Y(0) = Y - \psi T$$

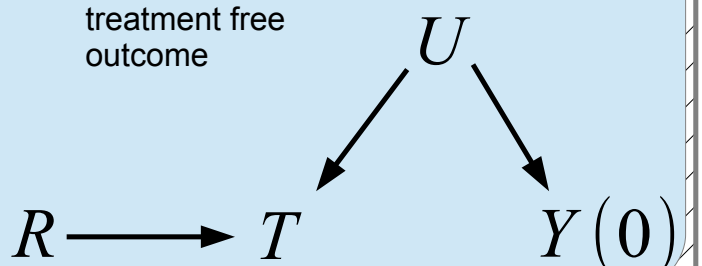
$$Cov(Y(0), R) = Cov(Y - \psi T, R) = Cov(Y, R) - \psi Cov(T, R) = 0$$

$$\psi = \frac{Cov(R, Y)}{Cov(R, T)} = \frac{ITT\ effect}{Complier\ fraction}$$

DAG for Observed outcome



DAG for treatment free outcome



Defiers can and do exist!

CACE in the presence of Defiers

Mean treatment arm outcome

$$P_1 = E[Y|R=1] = p_{at}\pi_{at} + p_{nt}\pi_{nt} + p_{1c}\pi_c + p_{1d}\pi_d$$

Mean control arm outcome

$$P_1 = E[Y|R=1] = p_{at}\pi_{at} + p_{nt}\pi_{nt} + p_{0c}\pi_c + p_{0d}\pi_d$$

$$\text{CACE estimate} = \frac{CACE\,\pi_c - DACE\,\pi_d}{\pi_c - \pi_d}$$

Testing the treatment effect homogeneity assumption

Using baseline covariates to identify two causal parameters and test the homogeneity assumption

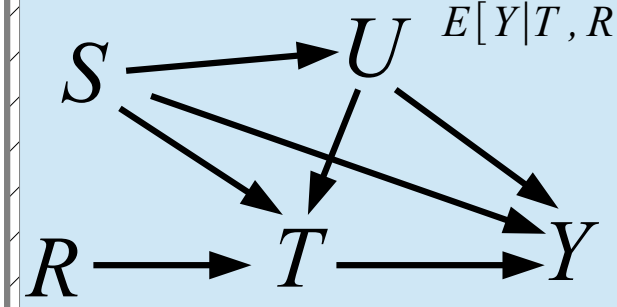
Treatment effect parameter

- ψ : Combined Treatment effect in Compliers and Always Takers
- ψ^* : Treatment effect in Always Takers only
- ψ_c : Treatment effect in Compliers only

Relationship

$$\psi = \frac{\psi_c \pi_c + \psi^* \pi_{at}}{\pi_c + \pi_{at}}$$

Estimating ψ and ψ^*



$$E[Y|T,R,S,U] = \beta_0 + \psi TR + \psi^* T(1-R) + \beta_s S + U$$ **Outcome model**

$$E[T|R,S] = \beta_0 + \beta_R R + \beta_S S + \beta_{SR} SR$$ **1st stage model**

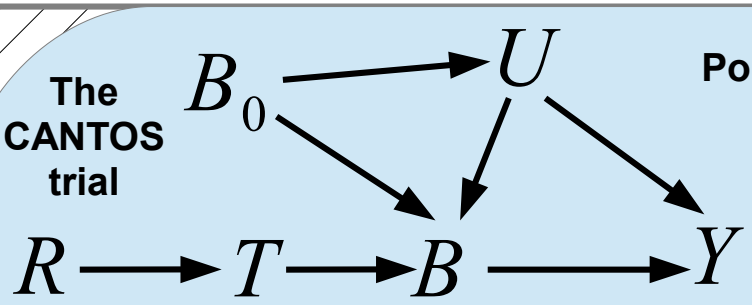
$$\hat{T}|R,S = \hat{\beta}_0 + \hat{\beta}_R R + \hat{\beta}_S S + \hat{\beta}_{SR} SR$$ **Fitted value**

$$E[Y|\hat{T},S] = \beta_0 + \psi \hat{T} R + \psi^* \hat{T}(1-R) + \beta_S S$$ **2nd stage model**

ITT estimator	Formula	Parameter form
ITT	$E[Y\|R=1] - E[Y\|R=0]$	$(\pi_c + \pi_{at})\psi - \psi^* \pi_{at}$
IV	ITT $/\pi_c$	$\psi_c = \frac{(\pi_c + \pi_{at})\psi - \psi^* \pi_{at}}{\pi_c}$
ITT in compliers	$E[Y\|R=1,T(1)=1,T(0)=0] - E[Y\|R=0,T(1)=1,T(0)=0]$	ψ
ITT in always takers	$E[Y\|R=1,T(1)=1,T(1)=0] - E[Y\|R=0,T(1)=1,T(1)=1]$	$\psi - \psi^*$

Deriving the Principal Stratum Estimand

The CANTOS trial



Policy Estimand $= E[Y|T=1] - E[Y|T=0]$

IV Estimand $= \dfrac{E[Y|T=1, B(1)=1, B(0)=0]}{E[Y|T=0, B(1)=1, B(0)=0]}$

$= \dfrac{\text{Policy Estimand}}{\text{Complier fraction}}$

Complier fraction $= Pr[B=1|T=1] - Pr[B=1|T=0]$

Principal Stratum Estimand $= E[Y|T=1, B(1)=1] - E[Y|T=0, B(1)=1]$

Derivation under treatment effect heterogeneity

Estimate ψ and ψ^* via interaction model

$$E[B|T, B_0] = \beta_0 + \beta_S B_0 + \beta_{B_0 T} B_0 T$$

$$\hat{B} = \hat{\beta}_0 + \hat{\beta}_{B_0} B_0 + \hat{\beta}_{B_0 T} B_0 T$$

$$E[Y|\hat{B}, T] = \beta_0 + \psi \hat{B} T + \psi^* \hat{B}(1-T)$$

Biomarker responder outcome mean in treatment group k (B(0) independent of T)

$$E[Y|T=k, B(1)=1] = \sum_{j=0}^{1} E[Y|T=k, B(1)=1, B(0)=j] Pr(B(0)=j), k=0,1$$

Biomarker responder outcome effect as a function Complier and Always responder causal effects

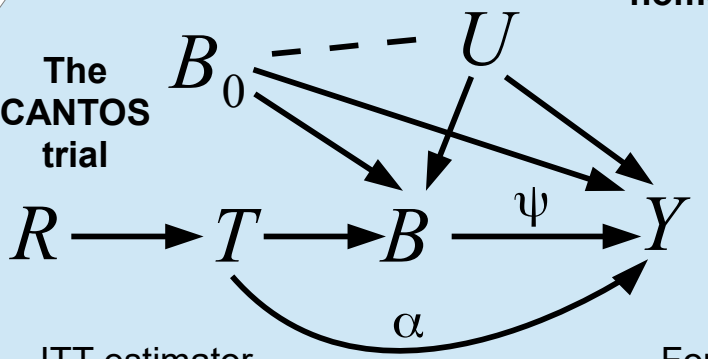
$$E[Y|T=1, B(1)=1] - E[Y|T=0, B(1)=1] = \psi Pr(B=0|T=0) + (\psi - \psi^*) Pr(B=1|T=0)$$

Principal Stratum estimand $= \psi - Pr(B=1|T=0)\psi^*$

Testing mechanism of action

Testing for direct and indirect effects of treatment under treatment effect homogeneity

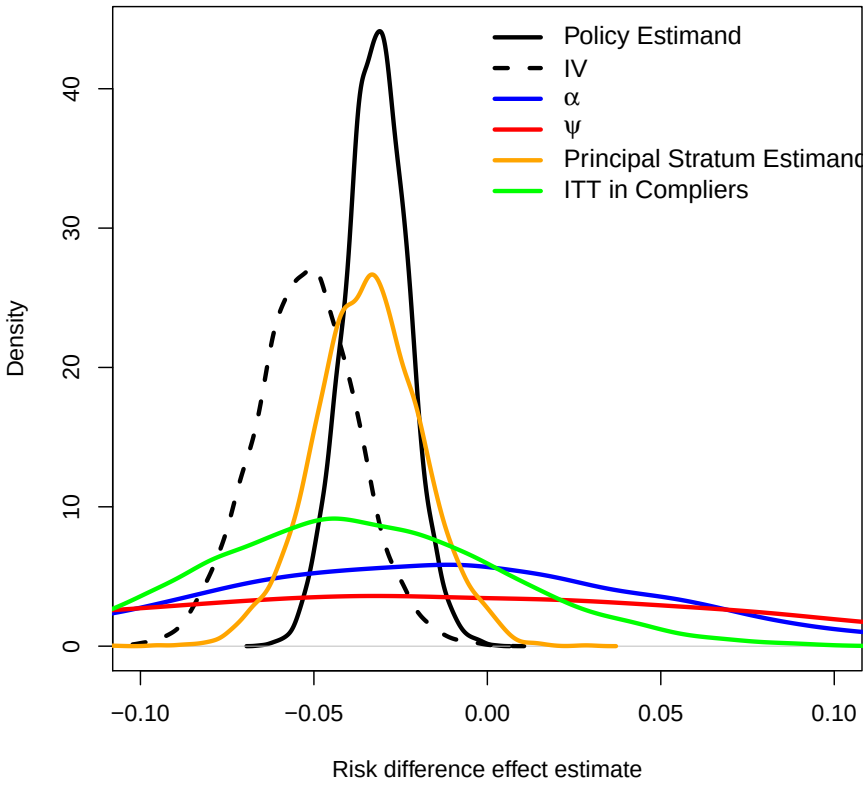
The CANTOS trial



Estimate ψ and α via

$$E[B|T, B_0] = \beta_0 + \beta_S B_0 + \beta_{B_0 T} B_0 T$$
$$\hat{B} = \hat{\beta}_0 + \hat{\beta}_{B_0} B_0 + \hat{\beta}_{B_0 T} B_0 T$$
$$E[Y|\hat{B}, T] = \beta_0 + \alpha T + \psi \hat{B} + \beta_{B_0} B_0$$

ITT estimator	Formula	Parameter
ITT (Policy)	$E[Y\|T=1] - E[Y\|T=0]$	$\alpha + \psi \pi_c$
IV	ITT $/\pi_c$	$\frac{\alpha}{\pi_c} + \psi$
ITT in compliers	$E[Y\|T=1, B(1)=1, B(0)=0] - E[Y\|T=0, B(1)=1, B(0)=0]$	$\alpha + \psi$
ITT in Always Responders	$E[Y\|T=1, B(1)=1, B(0)=1] - E[Y\|T=0, B(1)=1, B(0)=1]$	α
ITT in Treatment arm Responders, or the Principal Stratum Estimand	$E[Y\|T=1, B(1)=1] - E[Y\|T=0, B(1)=1]$	$\alpha + \psi Pr(B=0\|T=0)$



Estimand	Mean Estimate	Monte-carlo SD	Prop ≤ 0
Policy	-0.032	0.00906	1
IV	-0.0524	0.0148	1
α	-0.019	0.0691	0.614
ψ	-0.0214	0.112	0.578
Principal Stratum	-0.0344	0.0154	0.986
ITT in Compliers	-0.0403	0.0447	0.822

- Note how we can exploit -ve correlation again for ITT-in-complier and Principal Stratum Estimands!!