RTI $(h)(s)$

Health Solutions

September 27, 2018

Beyond Controlling for Confounding:
Design Strategies to Avoid Selection Bias and Improve Efficiency in Observational Studies

A Case Study of Screening Colonoscopy

Our Team

Xabier Garcia de Albeniz Martinez
+34 93.3624732
xgarcia@rti.org

Bradley Layton
919.541.8885
jblayton@rti.org

The power of **knowledge.**
The value of **understanding.**

Poll Questions

Poll question 1

Ice breaker: Which of the following study designs is best to evaluate the causal effect of a medical intervention?

❑ Cross-sectional study

❑ Case series

❑ Case-control study

❑ Prospective cohort study

❑ Randomized, controlled clinical trial

The power of **knowledge.**
The value of **understanding.**

We All Trust RCTs… Why?

- Obvious reasons

  - No confusion (i.e., exchangeability)

- Not so obvious reasons

  - Exposure represented at all levels of potential confounders (i.e., positivity)

  - Therapeutic intervention is well defined (i.e., consistency)

  - … And, because of the alignment of eligibility, exposure assignment and the start of follow-up (we'll soon see why is this important)

RCT = randomized clinical trial.

We just used the C-word: "causal" effect

Poll question 2

Which of the following is true?

- ❏ In pharmacoepidemiology, we work to ensure that drugs are effective and safe for the population

- ❏ In pharmacoepidemiology, we want to know if a drug **causes** an undesired toxicity

- ❏ Causal inference from observational data can be questionable, but being explicit about the **causal goal** and the validity conditions help inform a scientific discussion

- ❏ All of the above

In Pharmacoepidemiology, We Try to Infer Causes

ENCePP — European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

Official Title	Lead Investigator
Post-authorisation Safety Study: Risk of Out-of-Hospital Sudden Cardiac Death in Users of Domperidone, Users of Proton Pump Inhibitors, and Users of Metoclopramide	Dr Alejandro Arana
Cohort Study of the Relative Incidence of Major Cardiovascular Events Among Patients Initiating Prucalopride Versus a Matched Comparator Cohort	Dr Alicia Gilsenan

- These are good times to be an epidemiologist thanks to the wealth of data available (claims, electronic medical records, wearable devices, etc.)

- This helps facilitate precision and positivity… but the identifiability conditions have little to do with how large our database is

The power of **knowledge.**
The value of **understanding.**

What Are We Going to Talk About Today?

- Much of the focus in observational studies is placed on adjusting for confounding

  – This is **necessary** because we do not have control over the existence of common causes

- There's still room for bias when some basic design strategies are not followed

  – Be sure to align eligibility, exposure assignment and the start of follow-up

- Letting people contribute to your analysis whenever they are eligible by emulating a series of trials

Identifiability Conditions

To identify a causal effect, we need (i) data and (ii) assumptions external to the data

1. The conditional probability of receiving every value of treatment, though not decided by the investigators, depends only on the measured covariates (i.e., **conditional exchangeability**, no unmeasured confounding).

2. The values of treatment under comparison correspond to well-defined interventions that, in turn, correspond to the versions of treatment in the data (i.e., **consistency**)

3. The conditional probability of receiving every value of treatment is greater than zero (i.e., **positivity**)
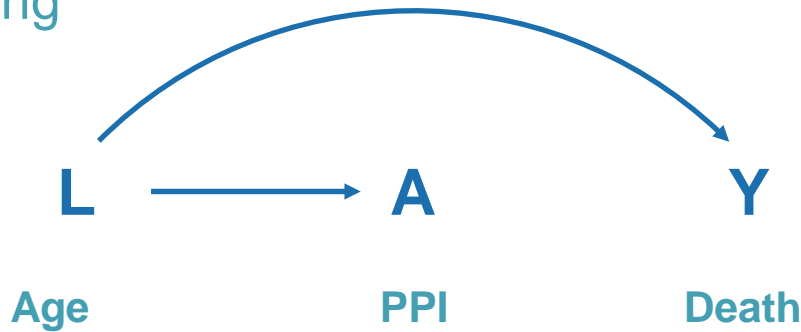
Hernán MA, Robins JM (2018). Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming.

The power of **knowledge.**
The value of **understanding.**

When Can the Identifiability Conditions Fail?

Identifiability Condition	RCT	Observational Analysis
Exchangeability	Losses to follow-up do not happen at random	We miss baseline confounders Losses to follow-up do not happen at random
Positivity	Artificial assignment of treatment guarantees it	Data are sparse or there are too many strata
Consistency	Protocol does not specify accurately the experimental intervention, or researchers do not follow it	Intervention is not well-defined or database does not differentiate multiple versions of the exposure (e.g., prevalent users)

Hernán MA, Robins JM (2018). Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming.

The power of **knowledge.**
The value of **understanding.**

# Why We (Think We) Love RCTs

Confounding

- Much of the attention in observational research is focused on managing confounding, which is fine

- How to handle confounding

  – <u>Measure the common causes</u> and use your favorite adjustment method

  – <u>Randomize</u>

PPI = proton pump inhibitor.

The power of **knowledge.**
The value of **understanding.**
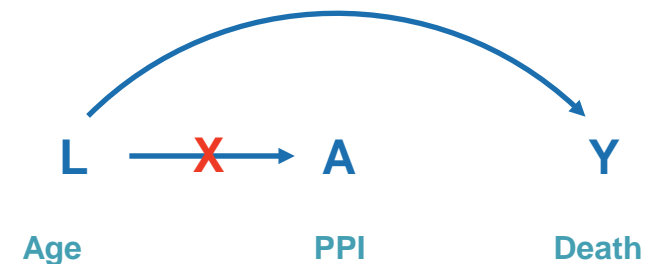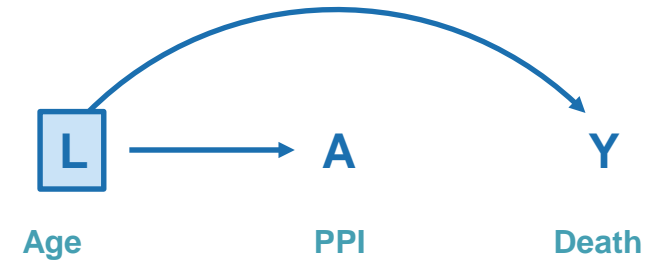
Let's Assume We Have Confounding and Positivity Under Control… What Else Can Go Wrong?

- Problems derived from the lack of synchronization in time of eligibility, treatment assignment, and time zero:
  - Time of eligibility (E): point in time when patients meet the eligibility criteria
  - Treatment assignment (A): point in time when patients are classified into exposure groups
  - Time zero (T_0): point in time when follow-up starts

- This is not a problem in RCT because of the following:
  - Time of eligibility: when deemed eligible by the PI
  - Treatment assignment: randomization happens shortly afterwards
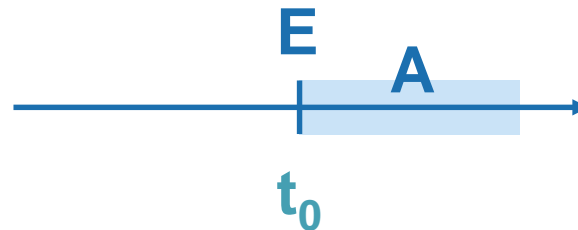 - Time zero: date of randomization.
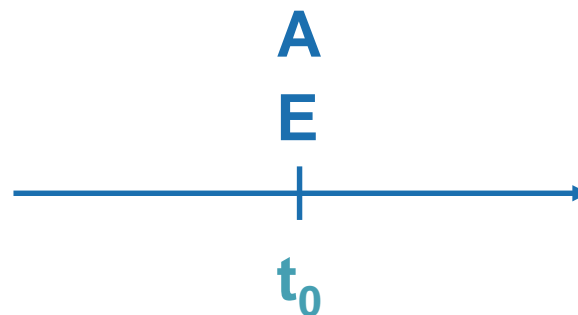
A

E

t_0

Is This a Problem in Observational Studies?

Some "EPI-101" biases are a consequence of this lack of synchronization

- **Classical immortal time bias**
  - Information on treatment after time zero is used to assign individuals to a treatment strategy. This time is immortal time.
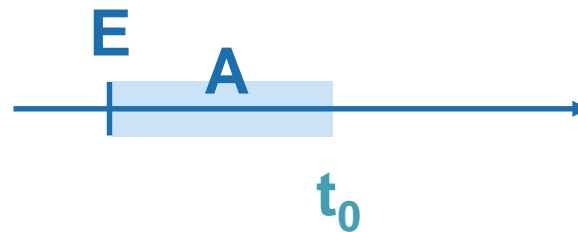

  - The definition of A by looking into the future guarantees that individuals are alive for that period of time

Is This a Problem in Observational Studies?

Some "EPI-101" biases are a consequence of this lack of synchronization

- **Prevalent/current user bias**
  - Inclusion of individuals who initiated the exposure of interest some time before the start of follow-up
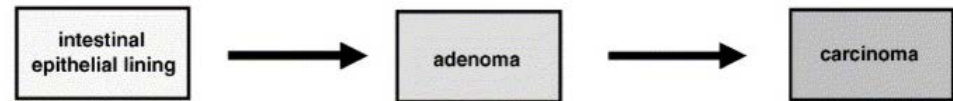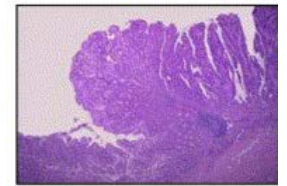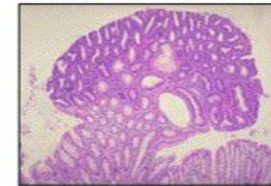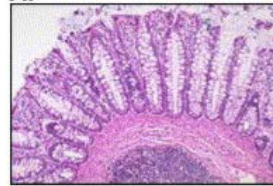

  - Prevalent users have survived the drug for a period of time
  - Past use of the drug can affect baseline covariates
  - Cannot inform health policy (i.e., cannot prescribe to be a "prevalent user")

Let's Continue the Conversation With a Case Study…

The power of **knowledge.**
The value of **understanding.**

CRC Screening: Intro

- CRC screening can prevent cancer

- CRC screening tools

  – Fecal occult blood test

  – Sigmoidoscopy

  – Colonoscopy

- RCTs have proved the following:

  – Sigmoidoscopy (either as a single intervention or twice in 3-5 years) reduces CRC incidence and CRC mortality

- No RCTs for colonoscopy (yet)

CRC = colorectal cancer.

The power of **knowledge.**
The value of **understanding.**
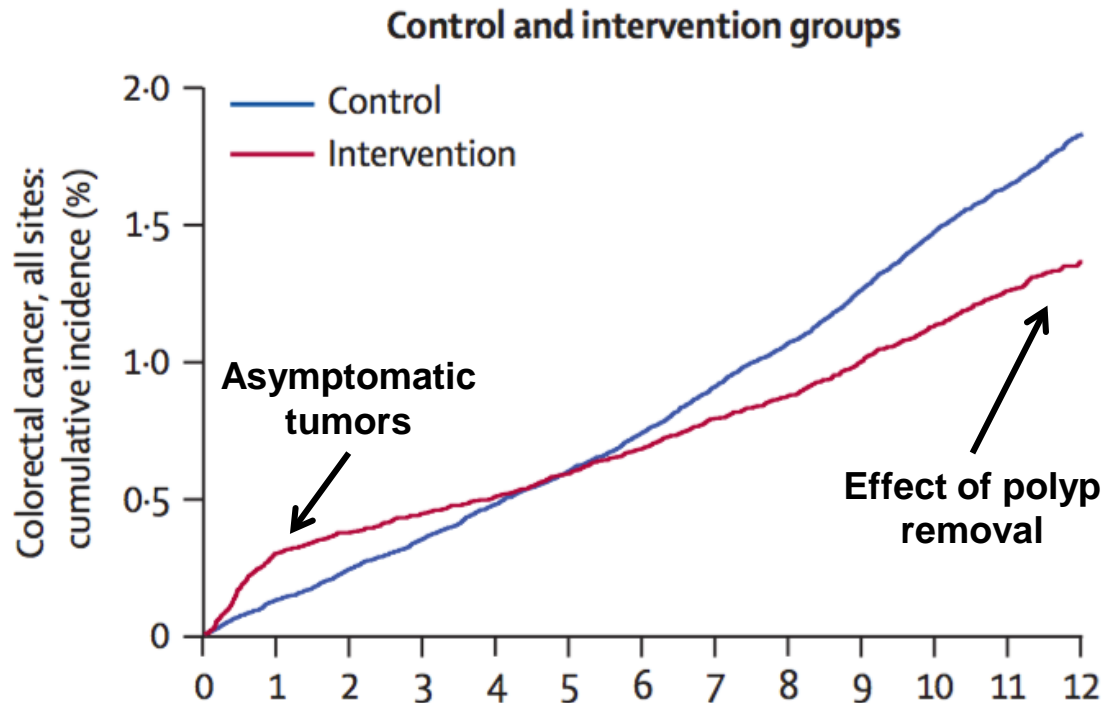
Research Question

- What is the effectiveness of screening colonoscopy in individuals aged 70-74?
  - Population barely (if at all) represented in ongoing RCTs
  - Over a decade of screening colonoscopy use in Medicare

- Let´s see
  - What challenges there are to answering this question using observational data (administrative data set)
  - How to deal with those challenges

Subtleties Specific to This Research Question

That we learned from sigmoidoscopy trials

- Screening sigmoidoscopy has very **little/no effect** on all-cause mortality:
  - RR = 0.98 (95% CI, 0.96-0.99)[1]

- The effect of screening sigmoidoscopy in CRC incidence is nonmonotonic.

Control and intervention groups

RR = risk ratio.

[1] Swartz AW, et al. Ann Intern Med. 2017v167:602–603. doi: 10.7326/M17-0859.

# Subtleties Specific to This Research Question

That we learned from sigmoidoscopy trials

- **All-cause mortality** is prone to be **confounded** more than CRC incidence in an observational setting

The power of **knowledge.**
The value of **understanding.**

RTI $(h)(s)$

Health Solutions

Once-Only Sigmoidoscopy in Colorectal Cancer Screening: Follow-up Findings of the Italian Randomized Controlled Trial—SCORE

Nereo Segnan, Paola Armaroli, Luigina Bonelli, Mauro Risio, Stefania Sciallero, Marco Zappa, Bruno Andreoni, Arrigo Arrigoni, Luigi Bisanti, Claudia Casella, Cristiano Crosta, Fabio Falcini, Franco Ferrero, Adriano Giacomin, Orietta Giuliani, Alessandra Santarelli, Carmen Beatriz Visioli, Roberto Zanetti, Wendy S. Atkin, Carlo Senore; and the SCORE Working Group
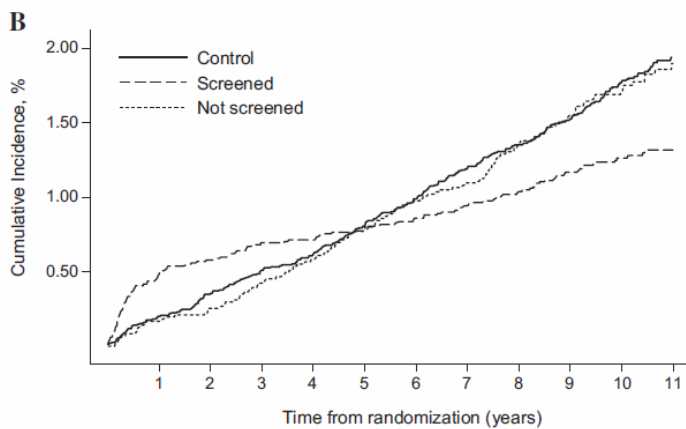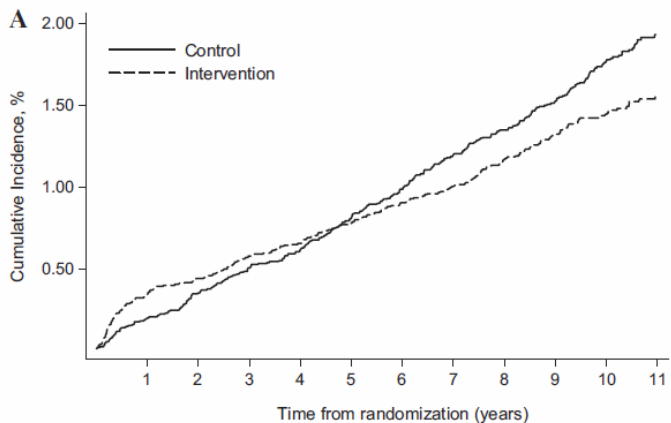
Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial

Wendy S Atkin, Rob Edwards, Ines Kralj-Hans, Kate Wooldrage, Andrew R Hart, John M A Northover, D Max Parkin, Jane Wardle, Stephen W Duffy, Jack Cuzick, UK Flexible Sigmoidoscopy Trial Investigators

The power of **knowledge.**
The value of **understanding.**
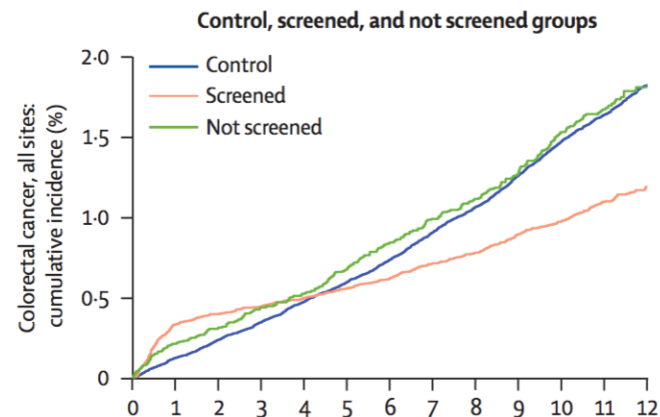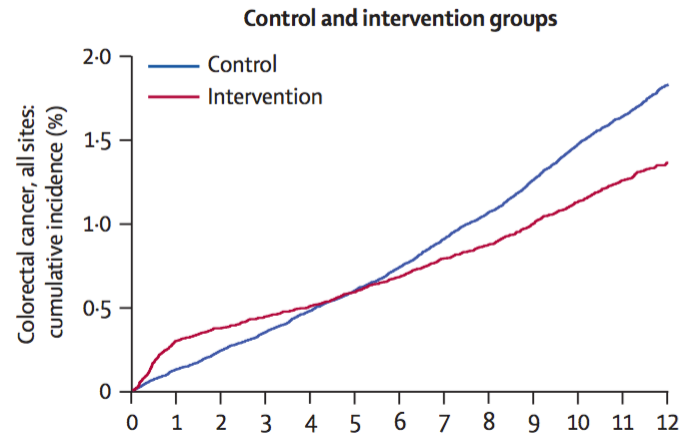
Original Investigation

Effect of Flexible Sigmoidoscopy Screening on Colorectal Cancer Incidence and Mortality
A Randomized Clinical Trial

Øyvind Holme, MD; Magnus Løberg, MD; Mette Kalager, MD, PhD; Michael Bretthauer, MD, PhD;
Miguel A. Hernán, MD, DrPH; Eline Aas, PhD; Tor J. Eide, MD, PhD; Eva Skovlund, PhD; Jørn Schneede, MD, PhD;
Kjell Magne Tveit, MD, PhD; Geir Hoff, MD, PhD

A Overall colorectal cancer incidence

Hazard ratio 0.80 (95% CI 0.70-0.92)
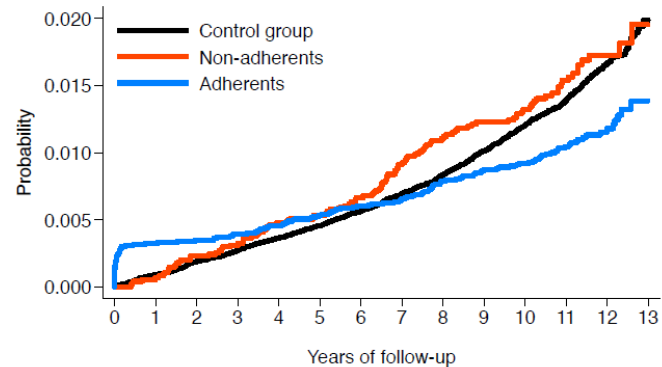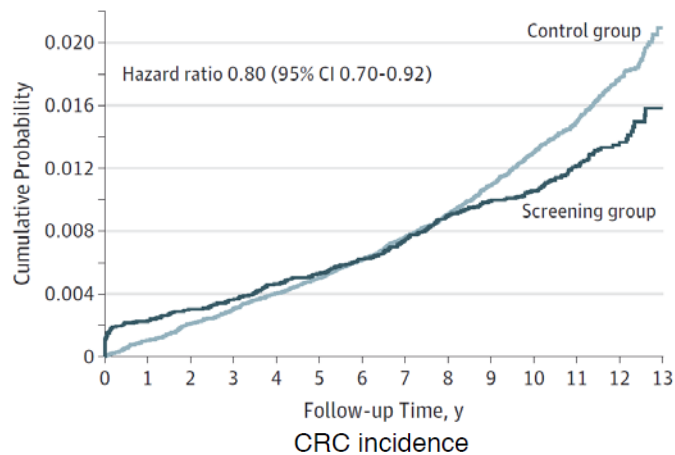
CRC incidence

All-cause mortality

HR: 0.97 (95% CI, 0.93-1.02)

HR = hazard ratio.

Subtleties Specific to This Research Question

Decisions

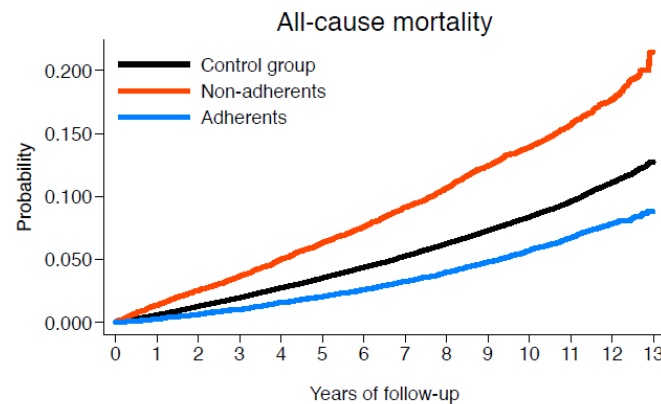- **All-cause mortality** is more prone to be **confounded** than CRC incidence in an observational setting
  - Stick to CRC incidence and stage at diagnosis
  - These are relevant clinical outcomes and plausibly the main mediators in improving cancer-specific survival

- The effect of screening colonoscopy in **CRC incidence is nonmonotonic**.
  - Use a cohort design to plot cumulative incidence curves
  - Standardize the cumulative incidence curves using a discrete-time hazards model[1]
  - Estimate the absolute difference at the end of the follow-up

[1] Hernán MA. The Hazards of Hazard Ratios. Epidemiology. 2010;21(1):13-15. doi:10.1097/EDE.0b013e3181c1ea43.

The power of **knowledge.**
The value of **understanding.**

First Step: Emulation of the Target Trial

Component	Target Trial	Emulated Trial Using Medicare
Aim	To estimate the effect of screening colonoscopy on the 8-year risk of CRC	Same
Eligibility	Individuals aged 70-74 No previous CRC, adenomas, IBD, or screening in previous 5 years	Same, plus continuous enrollment in Medicare
Treatment strategies	1. Screening colonoscopy at baseline 2. No screening for CRC at baseline	Same
Outcome	CRC diagnosis within 8 years	Same
Causal contrast	Intention-to-treat effect Per-protocol effect	Observational analog of a per-protocol effect

IBD = inflammatory bowel disease.

The power of **knowledge.**
The value of **understanding.**

Design No. 1

- We choose a **calendar date (e.g., January 1, 2004) as an arbitrary time zero** to start the follow-up and apply eligibility criteria

  - We assign eligible individuals to the colonoscopy arm if they received a colonoscopy in the previous 5 years or to the no colonoscopy arm otherwise
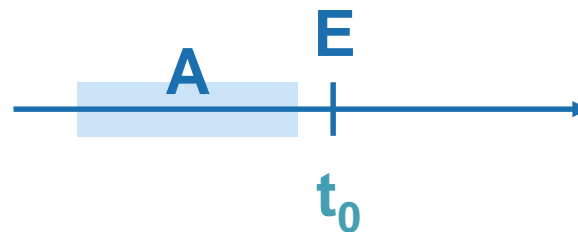
  - Colonoscopies performed before time zero can affect eligibility (e.g., by detecting adenomas, tumors, IBD). This is an example of selection bias.

(e.g., genetic predisposition)
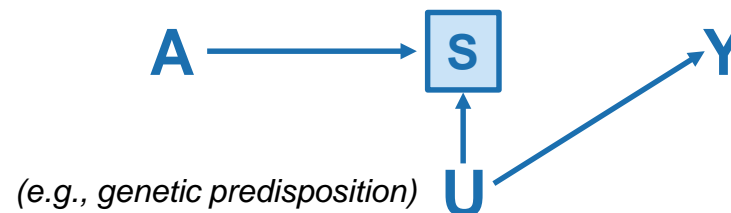
Let's Assume We Have Confounding Under Control: *What Else Can Go Wrong?*

Design No. 1

– What do we find when we run this analysis in Medicare (1999-2012)?

- **Screening looks implausibly beneficial** during the whole follow-up (we were expecting the detection of asymptomatic tumors at baseline)

Treatment before t_0, eligibility at t_0
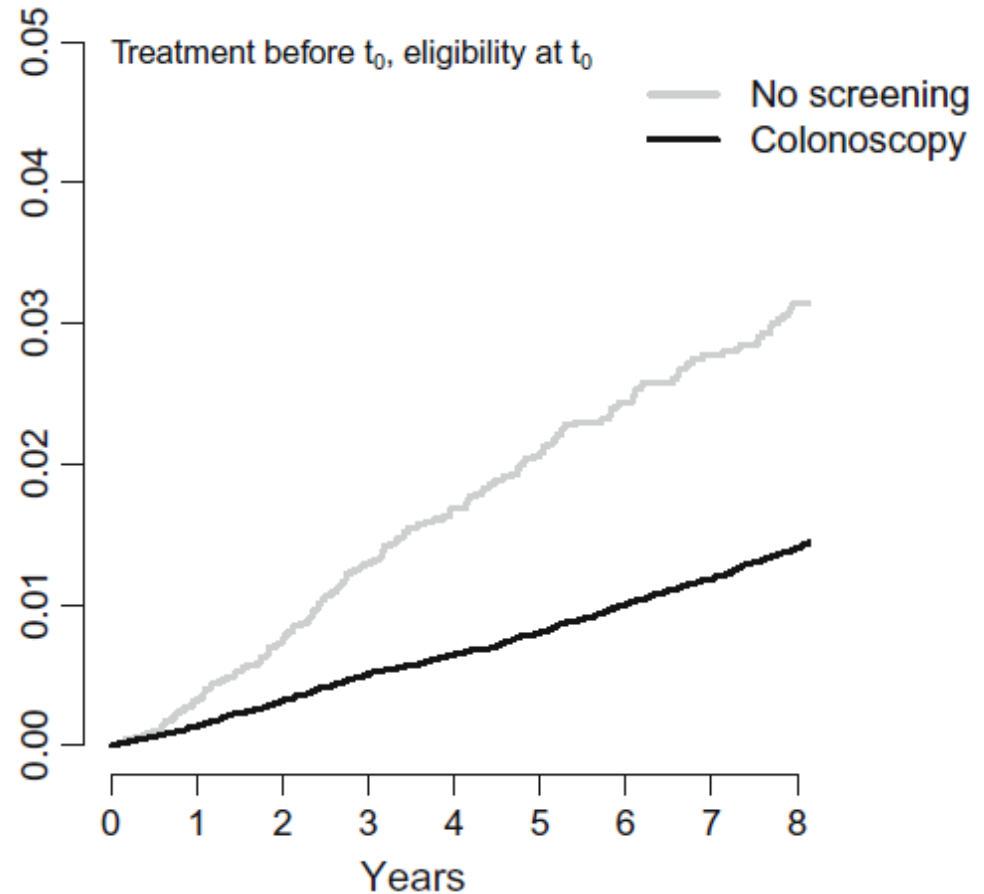
No screening
Colonoscopy

The power of **knowledge.**
The value of **understanding.**

Design No. 2

– Instead of choosing a calendar date as the anchor date:

- Colonoscopy arm: eligible individuals who receive a colonoscopy, t_0 = time of colonoscopy

- No screening arm: eligible individuals who do not receive a colonoscopy during the whole follow-up, t_0 being their first eligible time

$$E$$
$$A$$
$$t_0$$
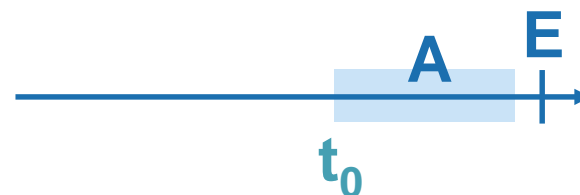
- Most of the CRCs are diagnosed with a colonoscopy, thus individuals in the no screening group do not have an opportunity to have a CRC diagnosed. This is another example of selection bias.

$$A \longrightarrow S \qquad Y^* \quad \textit{CRC diagnosis}$$

$$Y \qquad \textit{undiagnosed CRC}$$
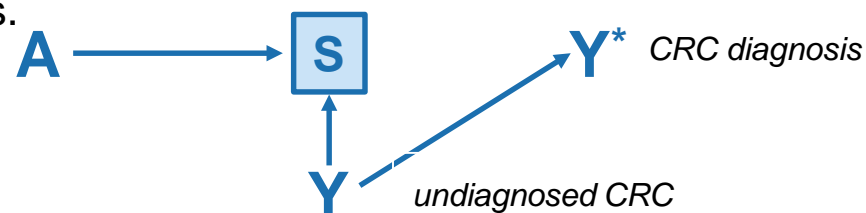
Let's Assume We Have Confounding Under Control: *What Else Can Go Wrong?*

Design No. 2

– **No screening looks implausibly beneficial** during the whole follow-up
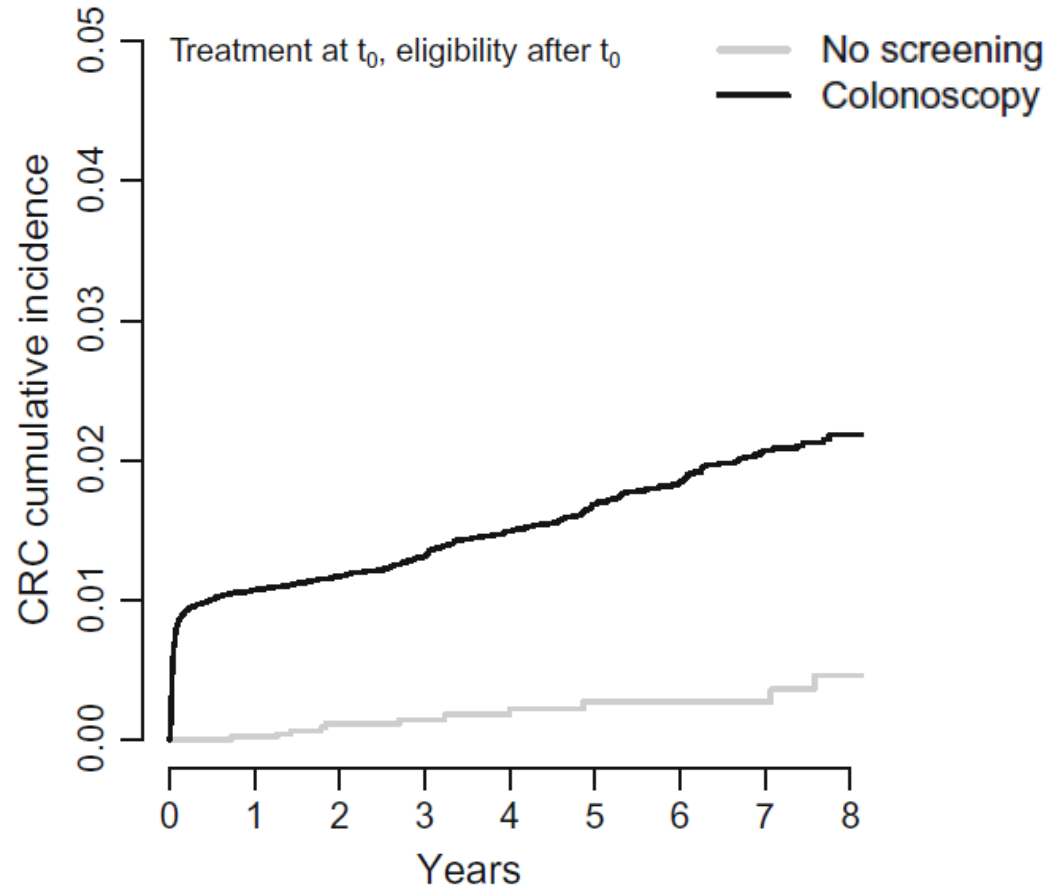
Let's Assume We Have Confounding Under Control: *What Else Can Go Wrong?*

Problems with the synchronization of the application of eligibility criteria, exposure assignment and the start of follow-up

Design No. 3

- Colonoscopy arm: eligible individuals who receive a colonoscopy, t_0 = time of colonoscopy.

- No screening arm: individuals who do not receive a colonoscopy at first eligibility time, t_0 = first eligible time.

A
E
_____|_____→
|
t_0

- This approach appropriately emulates the target trial, no selection bias

- If first eligibility happens earlier than the first colonoscopy, this can unbalance groups (e.g., younger individuals in the control group).

Problems with the synchronization of the application of eligibility criteria, exposure assignment and the start of follow-up

Design No. 3

– Graph more similar to sigmoidoscopy RCTs
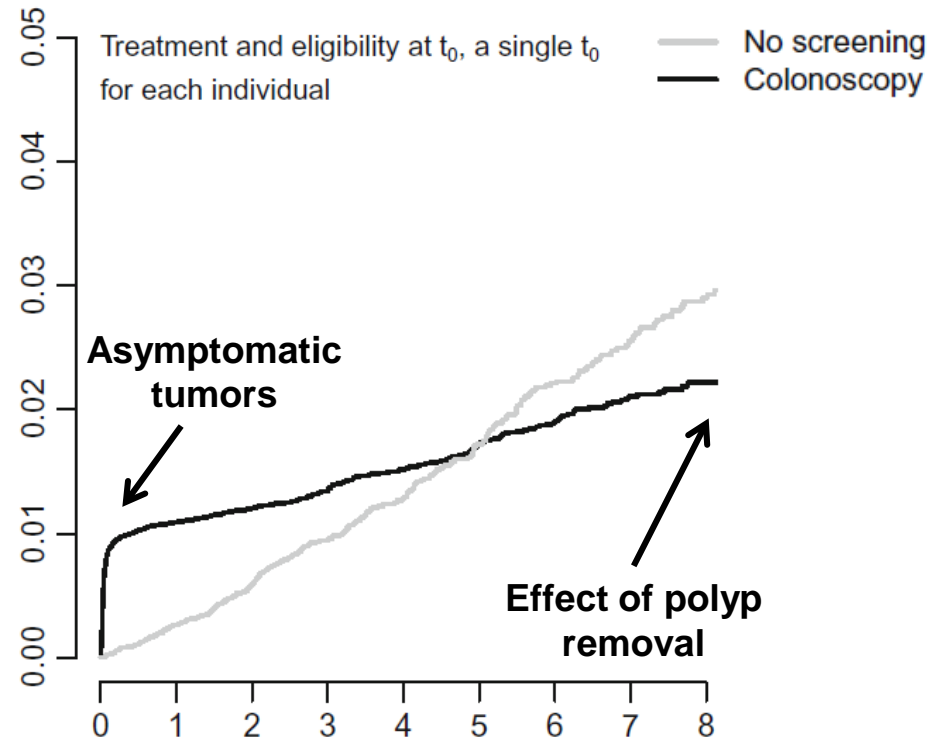
The power of **knowledge.**
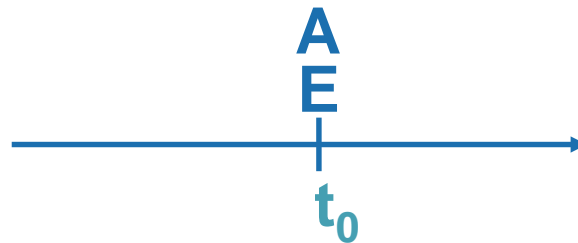The value of **understanding.**

Let's Assume We Have Confounding Under Control: *What Else Can Go Wrong?*

Problems with the synchronization of the application of eligibility criteria, exposure assignment and the start of follow-up

To summarize:

Design	Treatment Assigned	Eligibility Determined	Individuals Used More Than Once	Arm	N	CRC Cases	8 Year Risk Difference, % (95% CI)
1	Before t_0	At t_0	No	No screening	6,507	178	Ref.
				Screening	37,844	492	−1.7 (−2.2, −1.3)
2	At t_0	After t_0	No	No screening	6,241	11	Ref
				Screening	46,872	685	1.7 (1.4, 2.1)
3	At t_0	At t_0	No	No screening	72,249	1,086	Ref
				Screening	46,872	685	-0,67 (−1.03, −0.28)

Let's Assume We Have Confounding Under Control: *What Else Can Go Wrong?*

Problems with the synchronization of the application of eligibility criteria, exposure assignment and the start of follow-up

Design No. 4   The closest we can get to the emulation of a clinical trial

- We choose an anchor date (e.g., January 22, 2004) for eligibility and t_0; we synchronize exposure assignment with that date (this would be the equivalent of the randomization date)

  - Colonoscopy arm: individuals who receive a **colonoscopy** in the next 7 days of the anchor date

  - No screening arm: individuals who do **not** receive a colonoscopy in the next 7 days (increasing the number of days can take us back to situation No. 2)

A
E

t_0
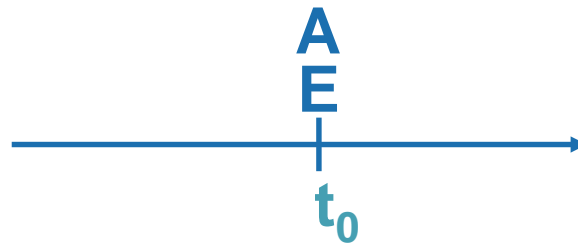
  - No selection bias

  - In our database, only 56 eligible individuals had a screening colonoscopy at that index date, with only 2 CRC diagnoses during the follow-up

  - These small numbers preclude the standardization of survival curves or obtaining precise estimates

Can We Make It Better?

Emulation of a sequence of target trials

- One way to increase efficiency is to emulate a sequence of target trials, starting at every interval during the study period. (i.e., same as situation No. 4 but at every available time interval)

- Let's say that this is our Medicare population, and we decide to implement an RCT at time interval 3

Time Interval

Legend:
- ⚪ Eligible person-time
- ⚪ Ineligible person-time because of no Medicare enrollment.
- ⊗ Ineligible person-time because of other exclusion criteria
- Ⓒ Eligible person-time, colonoscopy
- ⚫ Eligible person-time, outcome
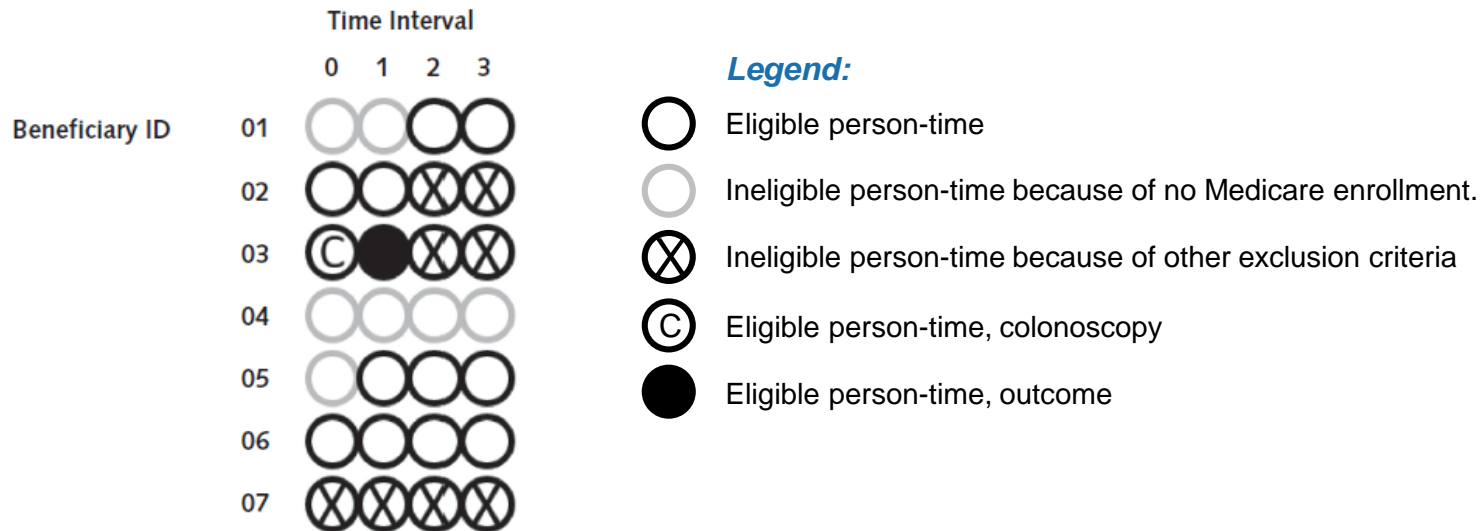
Can We Make It Better?

This is what we just did in situation No. 4

- We decide to implement an RCT at time interval 3

3 eligible individuals

Legend:

○ (bold) Eligible person-time

○ (light) Ineligible person-time because of no Medicare enrollment.

⊗ Ineligible person-time because of other exclusion criteria

Ⓒ Eligible person-time, colonoscopy

● Eligible person-time, outcome

Beneficiary ID	Group	Follow-up, Time Interval		Colorectal Cancer*	Baseline Covariate†
		Start	End		
01	Colonoscopy	3	4	0	1
05	No colonoscopy	3	5	1	0
06	No colonoscopy	3	6	0	1

ID = identification number.
* 0 = no; 1 = yes.
† For example, comorbidity score (ranges from 0 to 2).

The power of **knowledge.**
The value of **understanding.**

Can We Make It Better?

We can continue emulating trials over time to increase efficiency

- We decide to run an RCT at time interval 3… and then another one at time interval 4

Legend:

◯ Eligible person-time	ⓒ Eligible person-time, colonoscopy
◯ Ineligible person-time because of no Medicare enrollment.	⬤ Eligible person-time, outcome
⊗ Ineligible person-time because of other exclusion criteria	

The power of **knowledge.**
The value of **understanding.**

Can We Make It Better?

We can continue emulating trials over time to increase efficiency

- We decide to implement an RCT at time interval 3…
  and then another one at time interval 4… and at all time intervals

Beneficiary ID	Trial	Group	Follow-up, Time Interval		Colorectal Cancer*	Baseline Covariate†
			Start	End		
02	0	No colonoscopy	0	9	0	1
03	0	Colonoscopy	0	1	1	0
06	0	No colonoscopy	0	6	0	1
02	1	No colonoscopy	1	9	0	1
05	1	No colonoscopy	1	5	1	0
06	1	No colonoscopy	1	6	0	1
01	2	No colonoscopy	2	4	0	2
05	2	No colonoscopy	2	5	1	0
06	2	No colonoscopy	2	6	0	1
01	3	Colonoscopy	3	4	0	2
05	3	No colonoscopy	3	5	1	0
06	3	No colonoscopy	3	6	0	1
04	4	No colonoscopy	4	6	1	1
05	4	Colonoscopy	4	5	1	0
06	4	No colonoscopy	4	6	0	2
07	4	No colonoscopy	4	9	0	0
04	5	No colonoscopy	5	6	1	2
06	5	No colonoscopy	5	6	0	2
07	5	No colonoscopy	5	9	0	0
06	6	No colonoscopy	6	6	0	2
07	6	No colonoscopy	6	9	0	0
07	7	No colonoscopy	7	9	0	1
07	8	No colonoscopy	8	9	0	1
07	9	No colonoscopy	9	9	0	2

ID = identification number.
* 0 = no; 1 = yes.
† For example, comorbidity score (ranges from 0 to 2).

- Subtleties:
  - A single person can contribute to several trials (e.g., id 05 contributes to trials 1-4)
  - A single person can contribute to both arms (e.g., id 05 contributes to the "no colonoscopy" arm in trials 1-3 and to the "colonoscopy" arm in trial 4).
  - Baseline characteristics are extracted at each baseline (e.g., id 06 has a baseline value of 1 for trials 0-3 and a baseline value of 2 for trials 4-6)

The power of **knowledge.**
The value of **understanding.**

Can We Make It Better?

We can continue emulating trials over time to increase efficiency

- We decide to run an RCT at time interval 3… and then another one at time interval 4

ID 05 contributes to 2 cohorts and both arms
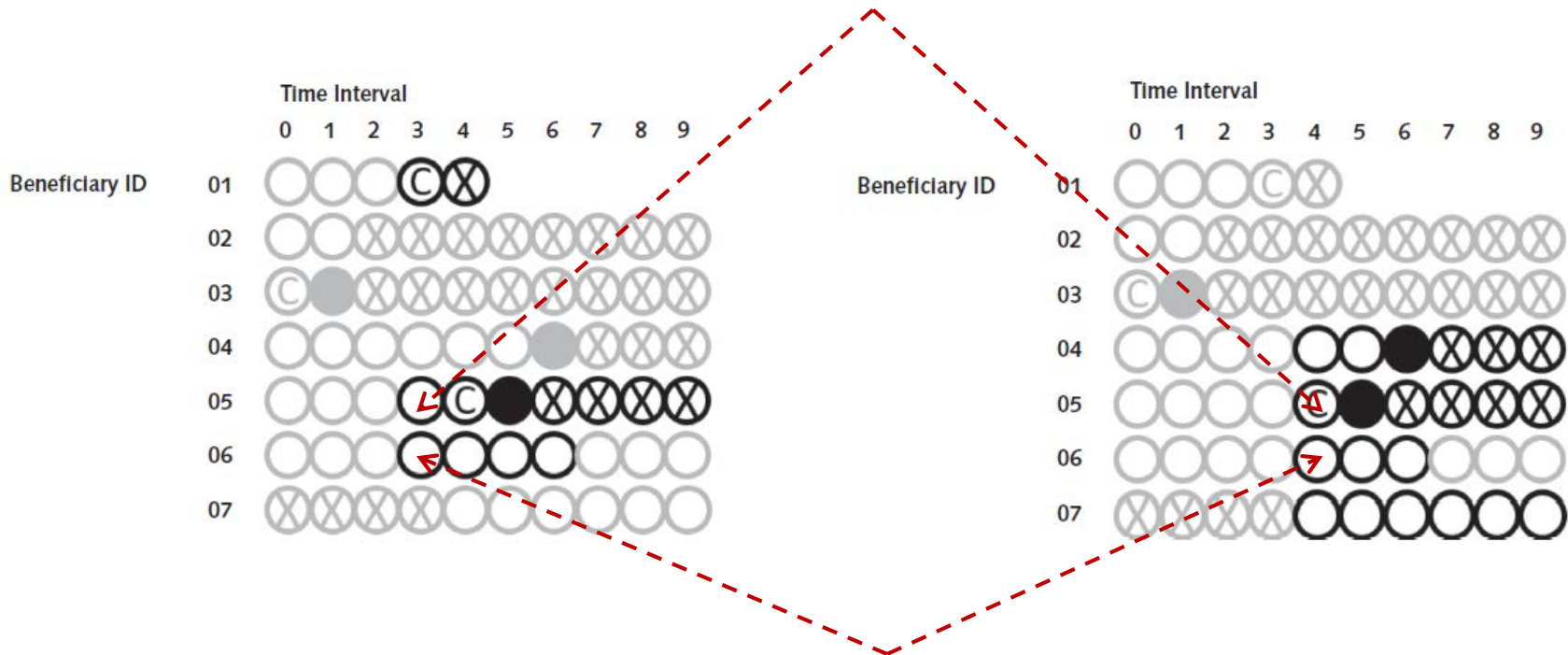
ID 06 contributes to 2 cohorts as a comparator

The power of **knowledge.**
The value of **understanding.**

What Is Gained With This Additional Complexity?

Precision

- In this specific scenario of Medicare claims (screening colonoscopy and CRC incidence), allowing the individuals to contribute to multiple emulated trials was the equivalent of increasing the sample size of the unexposed group approximately tenfold

Design	Treatment Assigned	Eligibility Determined	Individuals Used More Than Once	Arm	N	CRC Cases	Difference, % (95% CI)
1	Before t_0	At t_0	No	No screening	6,507	178	Ref.
				Screening	37,844	492	−1.7 (−2.2, −1.3)
2	At t_0	After t_0	No	No screening	6,241	11	Ref
				Screening	46,872	685	1.7 (1.4, 2.1)
3	At t_0	At t_0	No	No screening	72,249	1,086	Ref
				Screening	46,872	685	−0.67 (−1.03, −0.28)
4	At t_0	At t_0	Yes	No screening	1,762,816	21,954	Ref
				Screening	46,872	685	−0.63 (−0.83, −0.43)

What Is Gained With This Additional Complexity?
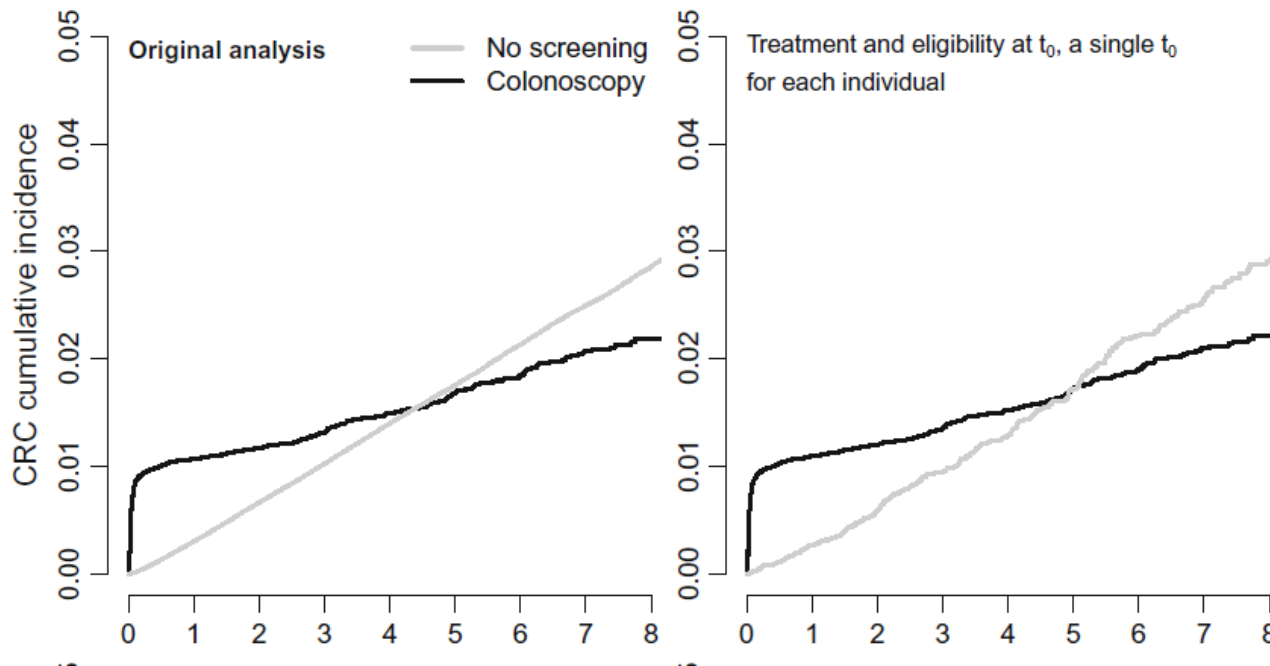
Precision

- In this specific scenario of Medicare claims (screening colonoscopy and CRC incidence), allowing the individuals to contribute to multiple emulated trials was the equivalent of increasing the sample size of the unexposed group approximately tenfold

The power of **knowledge.**
The value of **understanding.**

Take-Home Messages

- Confounding happens, we do not have control over it
  - Much of the focus in observational studies is placed on adjusting for confounding

- Selection bias can be a self-inflicted injury
  - Align eligibility, exposure assignment and the start of follow-up

- Let units of observation contribute to your analysis whenever they are eligible by emulating a series of trials to increase efficiency

The power of **knowledge.**
The value of **understanding.**

Further Reading

Methods

- Garcia-Albeniz, X, et al. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. Eur J Epidemiol. 2017;32(6): 495-500.

- Garcia-Albeniz X, Hsu J, Bretthauer M, Hernan MA. Effectiveness of Screening colonoscopy to prevent colorectal cancer among medicare beneficiaries aged 70 to 79 years: a prospective observational study. Ann Intern Med. 2017;166(1):18-26.

- Danaei G, Rodriguez LA, Cantero OF, Logan R, Hernan MA. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. Stat Methods Res. 2013;22(1):70-96.

- Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol. 2016;79:70-5.

- Hernan MA, Robins JM. Using Big data to emulate A target trial when A randomized trial is not available. Am J Epidemiol. 2016;183(8):758-64.

- Hernan MA, Robins JM, Garcia Rodriguez LA. Discussion on "statistical issues in the women's health initiative." Biometrics. 2005;61:922-30.

Applications

- Garcia-Albeniz X, Hsu J, Lipsitch M, Bretthauer M, Logan RW, Hernandez-Diaz S, et al. Colonoscopy and risk of infective endocarditis in the elderly. J Am Coll Cardiology. 2016;68(5):570-1.

- Danaei G, Garcia Rodriguez LA, Cantero OF, Logan RW, Hernan MA. Electronic medical records can be used to emulate target trials of sustained treatment strategies. J Clin Epidemiol. 2018;96:12-22.

- Hernan MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology (Cambridge, Mass). 2008;19(6):766-79.

- Schmidt M, Sorensen HT, Pedersen L. Diclofenac use and cardiovascular risks: series of nationwide cohort studies. BMJ (Clinical research ed). 2018;362:k3426.
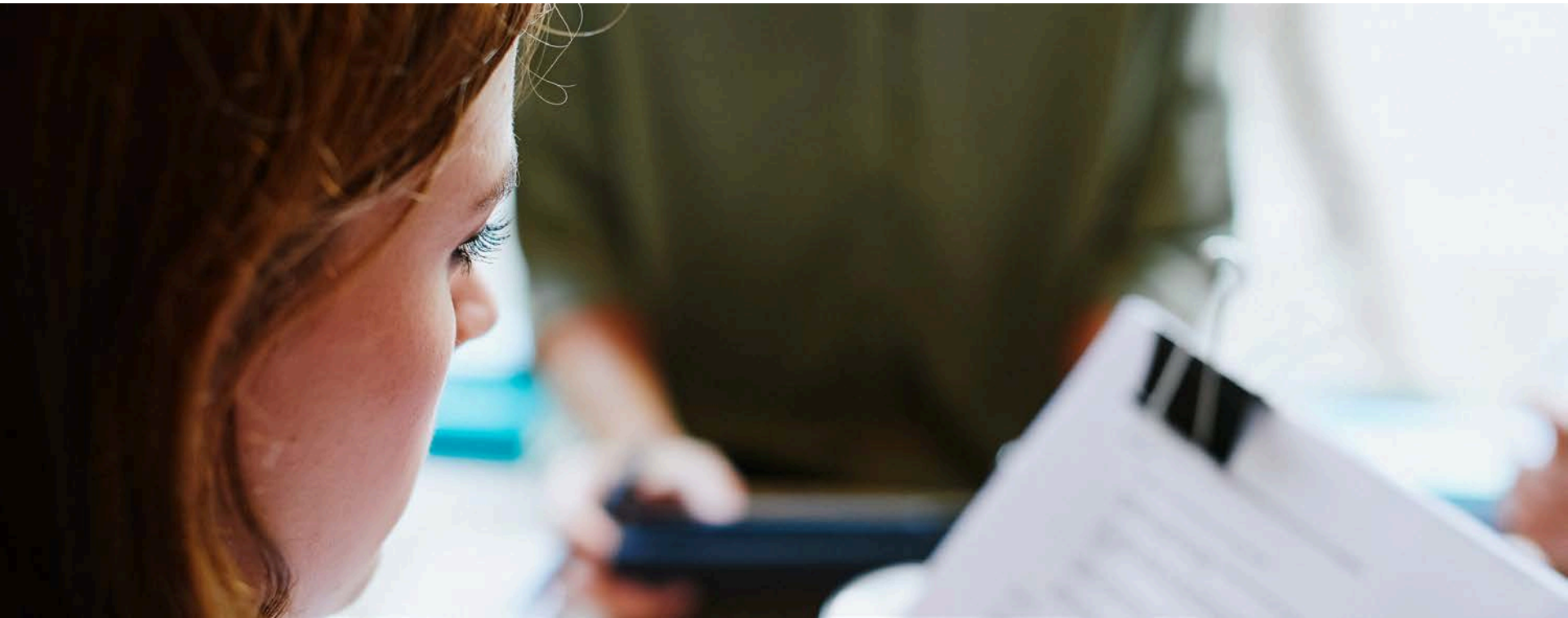
The power of **knowledge.**
The value of **understanding.**

Thank You
Questions?

The power of **knowledge.**
The value of **understanding.**

Generating knowledge and providing greater understanding so that you—and those who regulate, pay for, prescribe, and use your products—can make better decisions.

rtihs.org

The power of **knowledge.**
The value of **understanding.**

Contact Us

Xabier Garcia de Albeniz Martinez
+34 93.3624732
xgarcia@rti.org

Bradley Layton
919.541.8885
jblayton@rti.org

The power of **knowledge.**
The value of **understanding.**