How do we learn what works?

A two-step algorithm for causal inference from real world data

Miguel Hernán

DEPARTMENTS OF EPIDEMIOLOGY AND BIOSTATISTICS

HARVARD T.H. CHAN
SCHOOL OF PUBLIC HEALTH

@_MiguelHernan

1

Why do we want to know "what works"? Because decisions must be made NOW

- For clinical practice
  - Treat with A or with B?
  - Treat now or later?
  - Treat all individuals?
  - Stop all treatment?
- For public health
  - Implement a screening program?
  - At what age?
  - With what frequency?
  - Until what age?
- Decision making needs to be informed by causal knowledge about comparative effectiveness
  - and safety

2

How do we learn what works and what harms?
(How do we estimate causal effects?)

- The standard scientific answer:
  - Conduct a randomized experiment
- A relevant randomized trial would, in principle, answer each causal question about comparative effectiveness and safety
  - Interference/scaling up issues aside

3

But we rarely have randomized trials

expensive unethical impractical untimely




- And deferring decisions is not an option
  - no decision is a decision: “Keep status quo”
- What do we do?
  - We analyze observational data

4

Types of observational data

Research data

- Data collected specifically for research
  - Cohort studies, case-control studies, and other epidemiologic studies
  - Biobanks
  - Disease registries
  - ...

Found data

- Data generated for non-research purposes
  - Electronic health records
  - Insurance claims databases
  - National registers
  - ...

"Real world data"

"Routinely collected data"

5

We analyze observational data

because we cannot conduct a randomized trial

Observational analyses are **not** our preferred choice

- For each observational analysis for causal inference, we can imagine a hypothetical randomized trial that we would prefer to conduct
  - If only it were possible

6

The Target Trial

- The (hypothetical) randomized trial that we would like to conduct to answer a causal question
  - To learn what works and what harms
- A causal analysis of observational data can be viewed as an attempt to emulate some target trial
  - If we cannot translate our causal question into a target trial, then the question is not well-defined

7

The Target Trial

- Suggested more or less explicitly by many authors
  - Dorn (1953), Cochran, Rubin, Feinstein, Dawid…
  - for simple settings with a time-fixed treatment and a single eligibility point
- Explicit generalization to time-varying treatments and multiple eligibility points
  - Robins (1986)
  - Hernán, Robins. *Am J Epidemiol* 2016

8

The Target Trial concept leads to a simple algorithm for causal inference

1. Ask a causal question (point at the Target)
   - Specify the protocol of the Target Trial
2. Answer the causal question (shoot the Target)
   - Option A
     - Conduct the Target Trial
   - Option B
     - Use observational data to **explicitly** emulate the Target Trial
     - Apply appropriate causal inference analytics

9

Step 1 **Specify Target Trial protocol**	**Step 2** **Emulate Target Trial protocol**
☐ Eligibility criteria	☐ Eligibility criteria
☐ Treatment strategies	☐ Treatment strategies
☐ Randomized assignment	☐ Randomized assignment
☐ Start/End follow-up	☐ Start/End follow-up
☐ Outcomes	☐ Outcomes
☐ Causal contrast	☐ Causal contrast
☐ Analysis plan	☐ Analysis plan

10

Ok, so why is this a big deal?

- Why do we need to explicitly need to emulate a target trial when using observational data to learn what works?
- What happens if we just analyze the data as usual?
  - That is, if we compare "exposed" vs. "unexposed" and adjust for covariates?
- Let's see an example

11

EXAMPLE #1
Postmenopausal hormone therapy and heart disease

- Observational epidemiologic studies
  - >30% **lower risk** in current users vs. never users
    - e.g., hazard ratio: 0.68 in Nurses' Health Study
      - Grodstein et al. *J Women's Health* 2006
- Randomized trial
  - >20% **higher risk** in initiators vs. noninitiators
    - hazard ratio: 1.24 in Women's Health Initiative
      - Manson et al. *New England J Med* 2003

Shocking discrepancy!

12

The randomized trial
Women's Health Initiative (WHI)

- Double-blind
- Placebo-controlled
- Large
  - >16,000 U.S. women aged 50-79 yrs
- Randomly assigned to
  - estrogen plus progestin therapy
  - placebo
- Women followed approximately every year
  - for a maximum of 8 years

WHI randomized trial: Effect estimates

Intention-to-treat hazard ratio (95% CI) of coronary heart disase

- Overall 1.23 (0.99, 1.53)
- Years of follow-up
  - 0-2 1.51 (1.06, 2.14)
  - >2-5 1.31 (0.93, 1.83)
  - >5 0.67 (0.41, 1.09)
- Years since menopause
  - <10 0.89 (0.54, 1.44)
  - 10-20 1.24 (0.86, 1.80)
  - >20 1.65 (1.14, 2.40)

This hazard ratio can be fully explained by selection bias even if no woman benefits from hormone therapy
(Stensrud et al. *Epidemiology* 2017)

Why did observational studies get it “wrong”?

- Popular theory
  - Insufficient adjustment for lifestyle and socioeconomic indicators (residual confounding)
  - Corollary: causal inference from observational data is a hopeless undertaking
- An alternative theory
  - The observational studies were not emulating a target trial

15

WHI randomized trial compared women who **initiated** therapy with women who did not

- Design
  - Women randomly assigned to initiation of hormone therapy or placebo
  - Almost all women assigned to initiation received at least a dose, that is, they are classified as initiators
- Analysis
  - Compared risk between initiators (**incident** users) and noninitiators of hormone therapy
- This trial informs decisions about therapy initiation

16

Observational studies compared women **currently using** therapy with women who did not use it

- Design
  - Women were asked about therapy use
  - They were classified as current, past, or never users
- Analysis
  - Compared risk between current (**prevalent**) users and never users of hormone therapy
    - Was the estimate different from that of the WHI trial?
- What decision does this design/analysis inform?
  - What is the target trial?

17

"Current vs. never users" comparison is not clinically relevant

- Consider a woman wondering whether to start hormone therapy
  - The current vs. never contrast does not provide the information she needs
- Consider a woman wondering whether to stop hormone therapy
  - The current vs. never contrast does not provide the information she needs

18

What if we re-analyze the observational data…

… to explicitly emulate a target trial as close as possible to the WHI randomized trial?

- Causal inference algorithm
  - Step 1: Specify the protocol of a target trial of hormone therapy and coronary heart disease
  - Step 2: Emulate it
    - Hernán et al. *Biometrics* 2005; 61(4):922–930
    - Hernán et al. *Epidemiology* 2008; 19(6):766-779

19

Step 1 **Specify Target Trial protocol**	**Step 2** **Emulate Target Trial protocol**
Eligibility criteria	Eligibility criteria
Treatment strategies	Treatment strategies
Randomized assignment	Randomized assignment
Start/End follow-up	Start/End follow-up
Outcomes	Outcomes
Causal contrast	Causal contrast
Analysis plan	Analysis plan

20

Summary of Protocol of Target trial: Hormone therapy and coronary heart disease	
Eligibility criteria	Postmenopausal women with no history of cancer and other diseases, and no use of hormone therapy in the last 2 years.
Treatment strategies	1. Initiate estrogen plus progestin hormone therapy at baseline and remain on it during the follow-up, unless deep vein thrombosis, pulmonary embolism, myocardial infarction, or cancer are diagnosed 2. Refrain from taking hormone therapy during the follow-up
Assignment procedures	Participants will be randomly assigned to either strategy at baseline, and will be aware of the strategy they have been assigned to.
Follow-up period	Starts at randomization and ends at coronary heart disease diagnosis, death, loss to follow-up, or June 2000, whichever occurs earlier.
Outcome	Coronary heart disease diagnosed by a cardiologist
Causal contrasts	Intention-to-treat effect, per-protocol effect
Analysis plan	Intention-to-treat analysis, non-naïve per-protocol analysis

21

Important
Target trial must be a pragmatic trial

- Observational data cannot be used to emulate
  - a placebo-controlled trial
    - at most a trial with a “usual care” group
  - a trial with blind design
    - individuals are generally aware of the treatment they receive
  - treatment strategies that do not exist in the real world
  - enforcement of adherence to the protocol
  - tight monitoring that doesn’t happen in the real world

22

Observational data for emulation: The Nurses' Health Study

- Epidemiologic follow-up (cohort) study
- ~80,000 women with full data in 1980
- Information updated by questionnaire every two years
  - Use of hormone therapy
  - Diagnosis of coronary heart disease (confirmed by physician)
  - Medical diagnoses
  - Lifestyle data: diet, exercise, smoking…
  - Other risk factors for coronary heart disease

Emulation: Eligibility criteria

- Analysis is restricted to women who met the eligibility criteria of the target trial
  - Approximately equal to those of the WHI
- Including washout interval
  - no hormone use in 2-year period before baseline

Emulation: Treatment strategies

- Treatment strategies
  1) **Initiation** of oral estrogens plus progesterone therapy at baseline
  2) No hormone therapy **initiation** at baseline

- No blind assignment, no placebo control
  - Unlike the WHI randomized trial

Emulation: Outcome

- Identify women with a diagnosis of coronary heart disease during the follow-up
- Observational data cannot be generally used to emulate a target trial with systematic and blind outcome ascertainment
  - Except if outcome ascertainment cannot be affected by treatment history, e.g., if the outcome is mortality independently ascertained from a death registry

Emulation: Randomized assignment

- This is what “adjustment for confounding” means
- If insufficient data on confounders, then emulation of random assignment fails
  - Confounding bias
- Need to adjust for baseline covariates
  - via matching, stratification or regression, standardization or inverse probability (IP) weighting, g-estimation…

27

Emulation: Causal contrast

- Intention-to-treat effect: The effect of assignment to hormone therapy vs. no hormone therapy at baseline
  - regardless of future use during the follow-up
- We cannot consider the effect of assignment to hormone therapy at baseline
  - Because the dataset doesn’t include prescription dates
- We can consider the effect of initiation of hormone therapy at baseline
  - Analogous to a modified intention-to-treat approach in a trial
    - Including only those who take at least one dose of treatment

28

Emulation: Intention-to-treat analysis

- Compare risk between initiators and noninitiators of hormone therapy at baseline
  - regardless of future use during the follow-up
- Fit a Cox model (like the WHI did) with an indicator for treatment initiation and covariates
  - Age, past hormone use, parental history of myocardial infarction before age 60, education, husband's education, ethnicity, age at menopause, calendar time, high cholesterol, high blood pressure, diabetes, angina, stroke, coronary revascularization, osteoporosis, body mass index, smoking, aspirin use, alcohol intake, physical activity, diet score, multivitamin use, fruit/vegetable intake

Emulation summary

- We used the observational data to emulate a target trial with similar eligibility criteria, treatment arms, outcome, causal contrast, and analysis plan as the WHI randomized trial
- Some differences
  - Not blinded
  - Not placebo-controlled
  - Shorter average time since menopause than WHI
  - Longer follow-up than WHI

Effect estimates: hazard ratios (95% CIs)

	Randomized Women's Health Initiative	Observational Nurses' Health Study
☐ Overall	1.23 (0.99, 1.53)	1.05 (0.82, 1.34)
☐ Years of follow-up		
■ 0-2	1.51 (1.06, 2.14)	1.43 (0.92, 2.23)
■ >2	1.07 (0.81, 1.41)	0.91 (0.72, 1.16)
☐ Years since menopause		
■ <10	0.89 (0.54, 1.44)	0.88 (0.63, 1.21)
■ 10-20	1.24 (0.86, 1.80)	1.13 (0.85, 1.49)
■ >20	1.65 (1.14, 2.40)	--

31

When the target trial is explicitly emulated, then the same **causal question** is asked

- ☐ No shocking observational-randomized discrepancies
  - ■ though wide confidence intervals in both studies
- ☐ What about the popular hypothesis? Any residual confounding?
  - ■ Probably, but insufficient to explain the original discrepancy

32

Epidemiologic studies may be adequate to emulate target trials

- If high-quality observational data on treatment, outcome, and confounders are available
  - e.g., the Nurses' Health Study
- But most observational research relies on real world data
- Can emulation of a target trial work with large databases of real world data?
  - Let's see some examples

Examples of Target Trial emulation using different types of observational data

1. Hormone therapy and heart disease ← Research data: Epidemiologic study
2. Statins and mortality in cancer patients ← Research data: Cancer registry; Real world data: Insurance claims
3. Screening colonoscopy and cancer ← Real world data: Insurance claims
4. Statins and coronary heart disease ← Real world data: Electronic health records
5. Epoetin therapy and mortality in dialysis patients ← Real world data: Insurance claims + supplementary data
6. Antiretrovirals and mortality in HIV-positive individuals

EXAMPLE #2
Statins and mortality in cancer patients

- Statins are drugs that lower LDL-cholesterol
- In observational studies of cancer patients, statin use is associated with 30% lower mortality
  - Statins inhibit cancer growth?
- However, those studies did not attempt to explicitly emulate a target trial
- We did
  - Emilsson et al. *JAMA Oncology* 2018

35

Summary of Protocol of Target trial: Statin therapy and mortality in cancer patients	
Eligibility criteria	Individuals with Stage I-III colorectal, breast, prostate, and bladder cancer diagnosed at age 66 years or older, enrolled in Medicare parts A-B-D, and who did not receive a statin prescription in the previous 6 months.
Treatment strategies	1. Initiate statin therapy within 6 months of cancer diagnosis; discontinuation at any time that is clinically indicated 2. Refrain from using statin therapy during the follow-up
Assignment procedures	Participants will be randomly assigned to either strategy at baseline, and will be aware of the strategy they have been assigned to.
Follow-up period	Starts at randomization and ends at death, loss to follow-up, or December 2011, whichever occurs earlier.
Outcome	Cancer-specific mortality and all-cause mortality
Causal contrasts	Intention-to-treat effect, per-protocol effect
Analysis plan	Intention-to-treat analysis, non-naïve per-protocol analysis

36

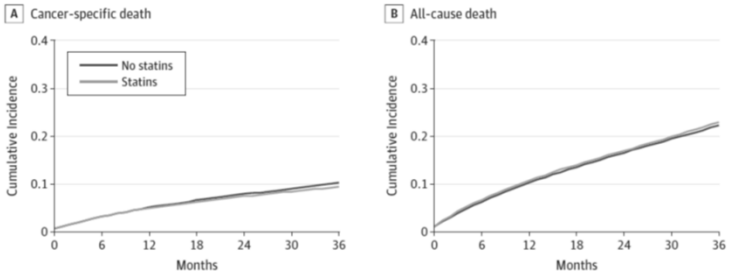
Observational data for emulation: SEER-Medicare

- SEER
  - cancer registries in 12 U.S. states
  - detailed information about cancer diagnosis
- U.S. Medicare
  - health insurance program for people 65 years or older (and others)
  - database includes insurance claims for all services provided, including statins, and death
- SEER-Medicare is the linkage of both

37

SEER-Medicare emulation: Hazard ratio estimates for statin vs. no statin initiation

- Cancer-specific mortality: 1.00 (0.88, 1.15)
- All-cause mortality: 1.07 (0.93, .21)



No beneficial effect of statins?
What about previous observational studies?

38

Selection bias in some observational studies

- Statin users at baseline vs. nonusers at baseline
  - Sounds familiar? No emulation of target trial

	Mortality hazard ratio (95% CI)	
	These studies	When we do that
Cancer-specific	0.77 (0.64, 0.89)	0.83 (0.76, 0.91)
All-cause	0.78 (0.67, 0.90)	0.83 (0.79, 0.87)

Immortal time bias in some observational studies

- Statin users at some point during the follow-up vs. nonusers during the follow-up
  - If you live longer, you are more likely to use statins

	Mortality hazard ratio (95% CI)	
	These studies	When we do that
Cancer-specific	0.35 (0.27, 0.44)	0.31 (0.24, 0.40)
All-cause	0.39 (0.33, 0.45)	0.57 (0.51, 0.63)

Emulating time zero (start of follow-up) is crucial to learn what works

- Criticisms of observational analyses often focus on residual confounding
  - failure to emulate randomization because of insufficient data on confounders
  - Hard to fix
- But many observational analyses have a more fundamental problem
  - Failure to choose time zero
  - Easy to fix

41

Step 1 **Specify Target Trial protocol**	**Step 2** **Emulate Target Trial protocol**
☐ Eligibility criteria	☐ Eligibility criteria
	☐ Treatment strategies
	☐ Randomized assignment
	☐ Start/End follow-up
☐ Outcomes	☐ Outcomes
☐ Causal contrast	☐ Causal contrast
☐ Analysis plan	☐ Analysis plan

Choosing time zero correctly: The low-hanging fruit for causal inference

42

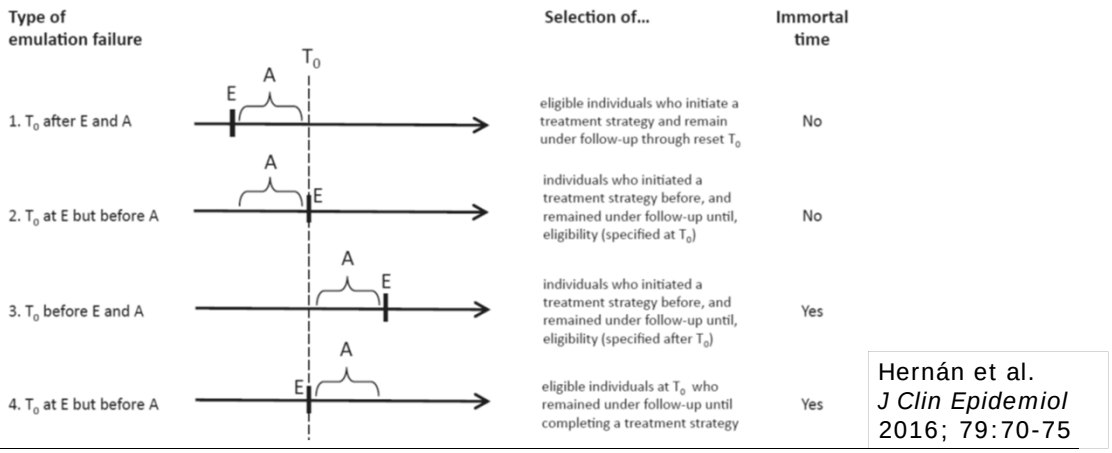
Time zero of follow-up in the Target Trial

- For each person, the time when 3 things happen
  - eligibility criteria are met
  - treatment strategies are assigned
  - study outcomes begin to be counted
- The same applies to observational analyses

- Misalignment of eligibility criteria and treatment assignment leads to selection bias / immortal time bias
  - Hernán et al. *J Clin Epidemiol* 2016; 79:70-75.

43

Misalignment of eligibility (E) and treatment assignment (A) prevents correct emulation



Hernán et al. *J Clin Epidemiol* 2016; 79:70-75

44

Why is it hard to align eligibility and treatment assignment at time zero?

- Time of eligibility may not be unique
  - An individual may meet the eligibility criteria at multiple times
- Treatment group may not be known at time zero
  - An individual's treatment strategy/exposure plan will be revealed after time zero

45

Emulation of time zero is not straightforward when there are multiple eligibility times

- In Example #2 (Statins in cancer patients), eligibility criteria are met as a single time
  - Cancer diagnosis
  - That's time zero
- In Example #1 (Hormone therapy), eligibility criteria may be met at different times
  - What's time zero?

46

Examples of Target Trial emulation using different types of observational data

1. Hormone therapy and heart disease ← Research data: Epidemiologic study
2. Statins and mortality in cancer patients ← Research data: Cancer registry; Real world data: Insurance claims
3. Screening colonoscopy and cancer ← Real world data: Insurance claims
4. Statins and coronary heart disease ← Real world data: Electronic health records
5. Epoetin therapy and mortality in dialysis patients ← Real world data: Insurance claims + supplementary data
6. Antiretrovirals and mortality in HIV-positive individuals

47

EXAMPLE #3
Screening colonoscopy and colorectal cancer

- Colonoscopy screening recommended at age 50 and then every 10 years in the U.S.
  - but its effectiveness never proven in randomized trials
  - 3 ongoing trials; results in 2025
- Very hard to conduct randomized trials
  - 10-15 years of follow-up are needed
  - >50,000 individuals needed
  - also, trials do not include older patients
- Need observational data to emulate a target trial

48

Summary of Protocol of Target trial Screening colonoscopy and colorectal cancer	
Eligibility criteria	Individuals aged 70–74 in 2004-2012 with no history of inflammatory bowel disease, adenoma, colectomy, and screening in the last 5 years; no gastrointestinal symptoms in last 6 months; continuous enrolment in Medicare for the last 5 years; at least 2 of the 3 preventive services offered yearly by Medicare (wellness visit, influenza vaccine, and breast or prostate cancer screening) in the previous 2 years
Treatment strategies	1. Screening colonoscopy at baseline 2. No screening colonoscopy at baseline
Assignment procedures	Participants will be randomly assigned to either strategy at baseline, and will be aware of the strategy they have been assigned to.
Follow-up period	Starts at randomization and ends at diagnosis of colorectal cancer, death, loss to follow-up, or January 2007, whichever occurs earlier.
Outcome	Colorectal cancer
Causal contrasts	Intention-to-treat effect, per-protocol effect
Analysis plan	Intention-to-treat analysis, non-naïve per-protocol analysis

49

The observational data: U.S. Medicare

- Federal health insurance program for people 65 years or older, with disabilities or with ESRD
  - About 50 million enrollees per year
- Random sample of Medicare claims dataset, 1999-2012
  - outpatient and inpatient services
  - doctor services
  - drug prescriptions
  - screening colonoscopy since July 2001
- Medicare enrollees can meet eligibility criteria at multiple times
  - every day since they turn 70 until 74

50

Choosing Time Zero when individuals meet eligibility at multiple times

Two unbiased choices:

- Choose a **single eligible time**
  - e.g., the first eligible time or a random eligible time
- Choose **every eligible time**
  - i.e., emulate a new trial starting at each eligible time
  - What we did for postmenopausal hormone therapy

Let's do both for colonoscopy screening

51

Choosing a single eligibility time as time zero

Garcia-Albeniz et al. *Eur J Epid* 2017

1. Colonoscopy group: individuals who meet the eligibility criteria and receive a colonoscopy
   - time zero is the time of the colonoscopy
2. No colonoscopy group: individuals who meet the eligibility criteria and did not receive a colonoscopy at first eligibility
   - time zero is, say, their first eligible time

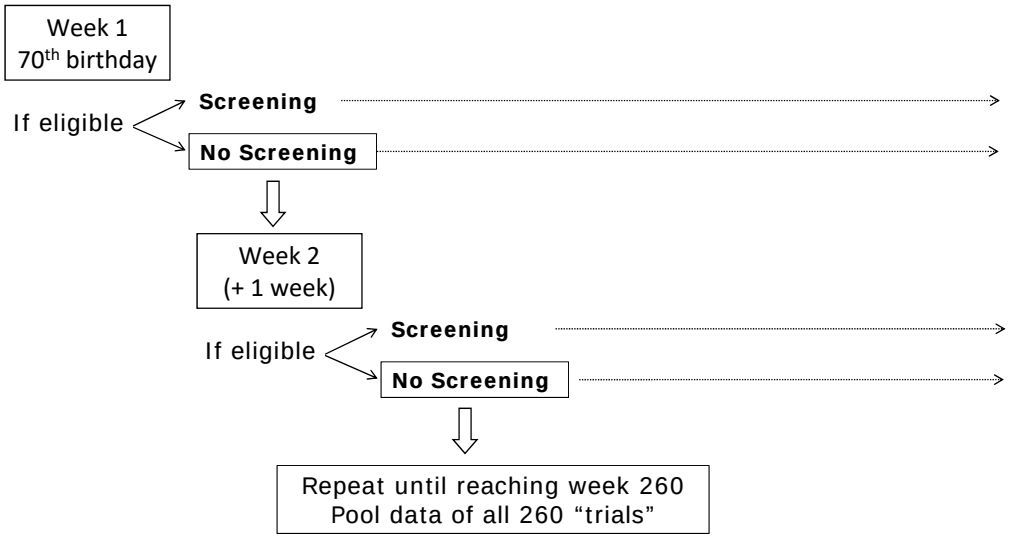
52

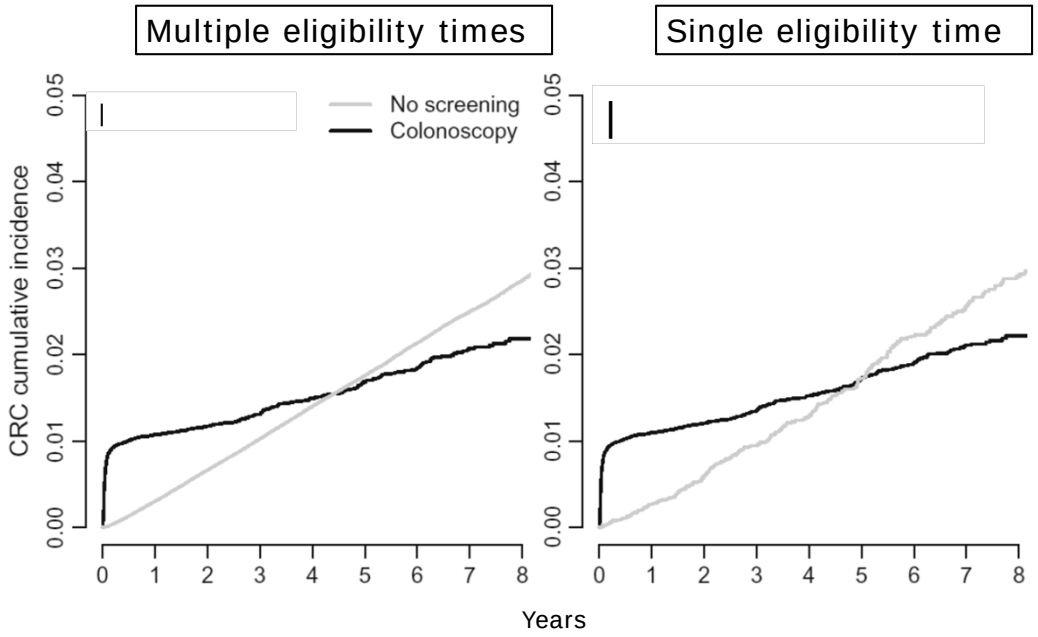
Choosing all eligible times as time zero

Garcia-Albeniz et al. *Ann Int Med* 2017; 166(1):18-26

- Emulate a new target trial each week of follow-up
  - Time zero is different in each trial
- Include in the emulation of each trial all individuals who are eligible at its corresponding time zero
- Combine all target trials for a more precise estimation
  - Need to take into account that some individuals will contribute to the emulation of several trials
    - Use a robust variance

Target trial: sequential emulation





Hernán - Target trial 55

55

Both approaches are valid choices of time zero

- Because they respect the basic principle of study design
  - Time zero is the time when eligibility is met and treatment strategies are assigned
- Consider two alternative observational analyses that do not respect this principle
  - and therefore do not emulate a target trial

Hernán - Target trial 56

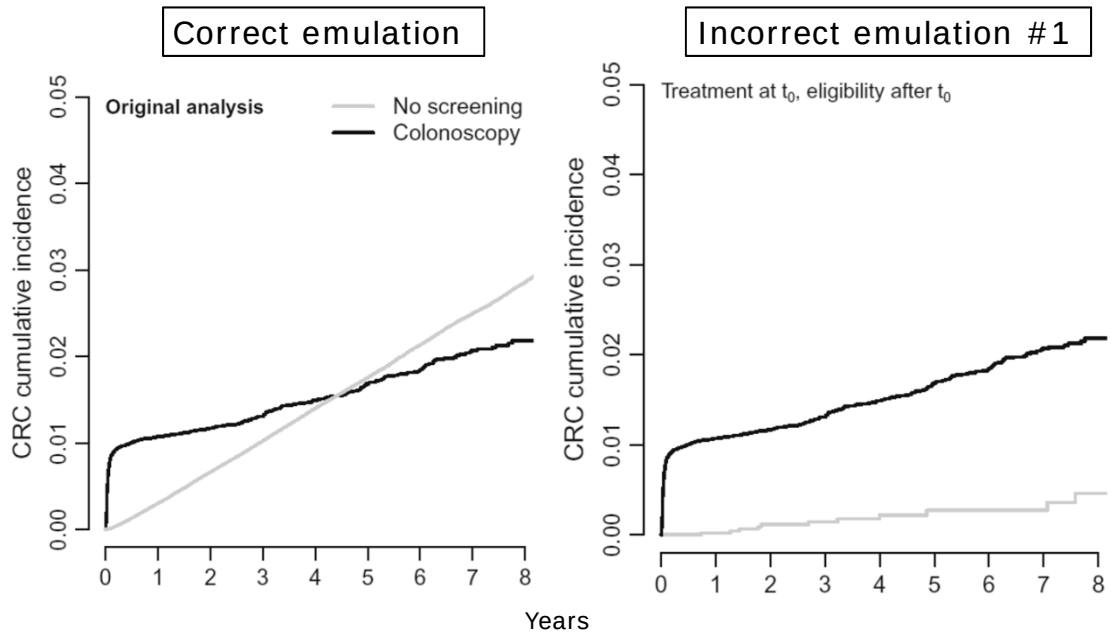
56

Incorrect emulation #1
Redefine the "No colonoscopy" group

1. Colonoscopy group: individuals who meet the eligibility criteria and receive a colonoscopy
   - time zero is the time of the colonoscopy

2. No colonoscopy group: individuals who meet the eligibility criteria and did not receive a colonoscopy ~~at first eligibility~~ during the follow-up
   - time zero is their first eligible time

57



58

Incorrect emulation #1
Biased

- Because most colorectal cancers are eventually diagnosed via colonoscopy
  - individuals in the no-screening group have little opportunity to be diagnosed
  - similar to naïve per-protocol analyses in randomized trials

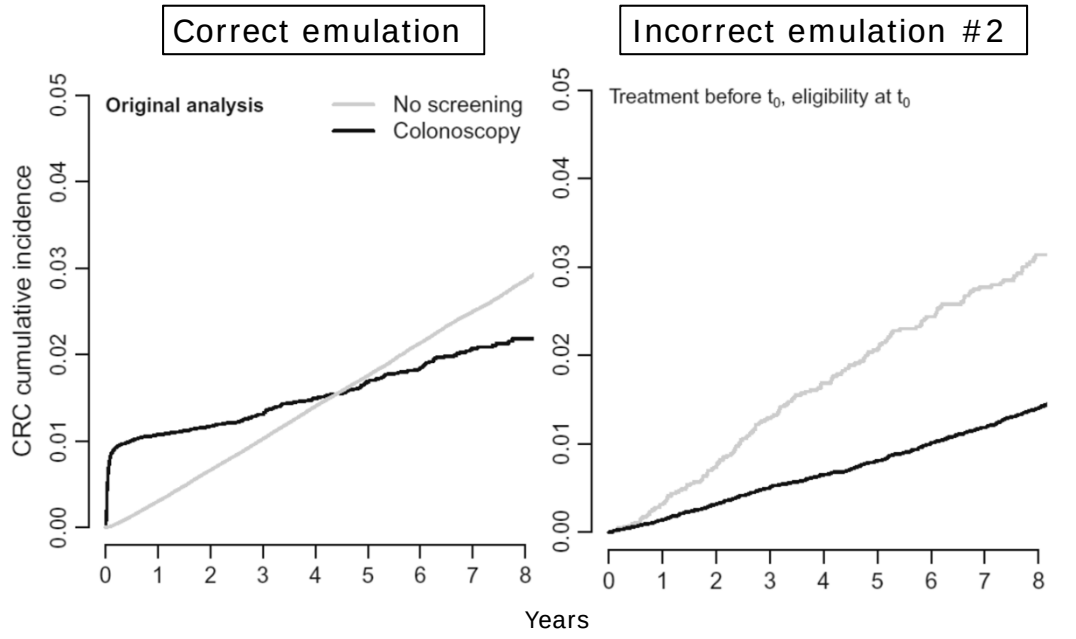
59

Incorrect emulation #2
Select arbitrary time zero and look back

1. Colonoscopy group: individuals who meet the eligibility criteria and received a colonoscopy in the five years before time zero
   - time zero is, say, January 2010

2. No colonoscopy group: individuals who meet the eligibility criteria and did not receive a colonoscopy in the five years before time zero
   - time zero is, say, January 2010

60



Incorrect emulation #2
Biased

- Because colonoscopies performed before assessing eligibility may affect eligibility
  - a colonoscopy that detects CRC or precursor lesions in the previous five years will result in the individual being excluded from the analysis
  - similar to approach that created confusion about the effect of postmenopausal hormone therapy in observational studies

Basic principle of study design for causal inference

- Treatment assignment and the determination of eligibility occur simultaneously at time zero
- In our example, observational analyses that violated this principle yielded implausible estimates
- Good news: correct time zero determination is always possible

63

2 key components of the emulation of the target trial

1. Randomized assignment
   - Emulation requires adjustment for confounding
2. Specification of time zero
   - Time zero must be synchronized with determination of eligibility and assignment of treatment strategies

- Lack of randomization is usually blamed for the failings of observational analyses, but…
  - we have seen that incorrect specification of time zero is often the actual culprit
- Next we will focus on confounding

64

These examples show that successful emulation of a Target Trial requires

- High-quality data on treatment, outcome, and confounders
  - If possible, assessment of data accuracy
    - Validation studies to quantify misclassification
    - Internal consistency checks to detect problems
    - Cross-datasets comparisons to flag coding differences
- Knowledgeable users of the data
  - Time-varying clinical workflows, idiosyncratic coding practices, software versions…
    - e.g., what does a “coronary heart disease” code mean? Maybe used when a physician suspected the diagnosis and ordered a test?

65

The target trial is typically a compromise

- between the ideal trial we would really like to conduct and the trial we may reasonably emulate using the available data

- The 2-step algorithm is typically iterative
  - Specifying the protocol of the target trial requires detailed knowledge of the database
  - The target trial approach allows you to systematically articulate the tradeoffs that you are willing to accept
    - regarding eligibility criteria, treatment strategies, outcomes

66

Not explicitly describing our causal goal is like shooting without a target

AJPH PUBLIC HEALTH OF CONSEQUENCE

Am J Public Health. **2018**;108: 616–619

The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data

Causal inference is a core task of science. However, authors and editors often refrain from explicitly acknowledging the causal goal of research projects; they refer to causal effect estimates as associational estimates.

Miguel A. Hernán, MD, DrPH

See also Galea and Vaughan, p. 602; Begg and March, p. 620; Ahern, p. 621; Chiolero, p. 622; Glymour and Hamad, p. 623; Jones and Schooling, p. 624; and Hernán, p. 625.

You know the story: Dear author: Your observational Confusion then ensues at the most basic levels of the scientific process and, inevitably, errors are glass of red wine per day versus no alcohol drinking. For simplicity, disregard measurement error and

67

Every time someone presents observational estimates to estimate causal effects, ASK

"What is the target trial?"

- If they look puzzled, help them specify the target trial
- If no target trial can be identified, ask them to start over

68

Thank you

- For more info
  - Twitter: @_MiguelHernan
  - www.hsph.harvard.edu/miguel-hernan/
- Causal Inference book
  - Free online, google “causal inference book”
  - www.facebook.com/causalinference