Why Target Trial Emulation

DR MICHALIS KATSOULIS

ASSOCIATE PROFESSOR IN BIOMEDICAL STATISTICS

What works to improve human health?

- Researchers of causal inference in epidemiology and public health are interested in investigating whether a specific treatment/intervention is beneficial (or not) for a health outcome

- For obvious reasons, epidemiologists would trust experimentation for this purpose, as it happens with biology, chemistry, physics, engineering. The experimental studies they use are the Randomised Controlled Trials (RCTs)

- However, there are lots of observational data (e.g. EHR, cohorts) that can provide some indications. How can we appropriately use them for comparative effectiveness research?

Individual treatment effects

- Ideally, we would like to know if a particular intervention works at an individual level for each one of us. Of course, this is not feasible…
- In reality, we know that, e.g., John was an individual with obesity who underwent bariatric surgery at age 43 and did not develop cancer after 10 years
- However, we do not know what would have happened if he didn't undergo bariatric surgery

Individual treatment effects

Very few times we can have evidence for individual treatment effects e.g.

1. Doctor cracked his own knuckles on one hand only for 60 years to prove it doesn't cause arthritis
2. Trials on medication on right versus left eye on glaucoma

Even in these case, we have to be lenient with the definition of these effects (i.e what would have happened to the doctor if he cracked his own knuckles on the other hand?)

Potential outcomes and ATE

	Bariatric surgery (X)	Cancer (Y)	$Y^{X=0}$	$Y^{X=1}$
John	1	0	-	0
Maria	1	1	-	1
Charles	0	0	0	-
Kate	0	1	1	-

- $Y^{X=0}$ is the (counterfactual or potential) outcome of an individual, had this individual i been untreated
- $Y^{X=1}$ is the (counterfactual or potential) outcome of an individual, had this individual i been treated
- Average treatment effect (ATE) = $E[Y^{X=1}] - E[Y^{X=0}]$, i.e. the expected value (probability) of developing cancer had everybody been treated $E[Y^{X=1}]$ minus the probability of developing cancer had nobody been treated $E[Y^{X=0}]$

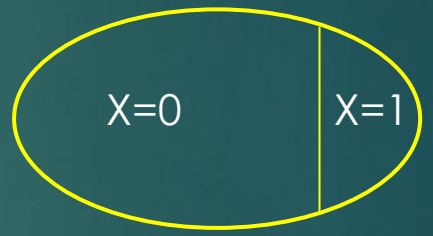
However, for our individuals, we don't know only what happened when they where both treated and untreated

Conditioning vs Potential outcomes

Example: We are interested in estimating the effect of bariatric surgery on the 10-year risk of cancer in individuals with BMI≥35kg/m^2

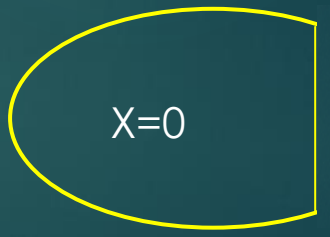
Exposure: X= Bariatric surgery (0: No, 1:yes)

Outcome: Y= cancer (0: No, 1: Yes)

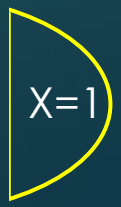


Notation for conditioning

E[Y=1 | X=0] is the expected value (probability) of developing cancer among individuals who are not treated



E[Y=1 | X=1] is the expected value (probability) of developing cancer among individuals who are treated



Conditioning vs Potential outcomes

Example: We are interested in estimating the effect of bariatric surgery on the 10-year risk of cancer in individuals with BMI≥35kg/m^2 from an RCT

Exposure: X= Bariatric surgery (0: No, 1:yes)

Outcome: Y= cancer (0: No , 1: Yes)

Notation for potential outcomes

$E[Y^{X=0}=1]$ or $E[Y^{0}=1]$ is the expected value (probability) of developing cancer had no individuals been untreated

X=0

$E[Y^{X=1}=1]$ or $E[Y^{1}=1]$ is the expected value (probability) of developing cancer had all individuals been treated

X=1

Estimand

	Bariatric surgery (X)	Cancer (Y)	$Y^{X=0}$	$Y^{X=1}$
John	1	0	-	0
Maria	1	1	-	1
Charles	0	0	0	-
Kate	0	1	1	-

Estimand: Quantity we want to estimate

In our case, we are interested in the effect of bariatric surgery on cancer, more specifically, we want to estimate the average treatment effect (ATE)

Using the potential outcome notation: $ATE = E[Y^{X=1}] - E[Y^{X=0}]$

In randomised trials $E[Y^{X=1}] = E[Y=1 \mid X=1]$ and $E[Y^{X=0}] = E[Y=1 \mid X=0]$

(Because of randomisation)

Estimand

In randomised trials, ATE can easily be calculated from the data

$$ATE = E[Y=1 \mid X=1] - E[Y=1 \mid X=0]$$

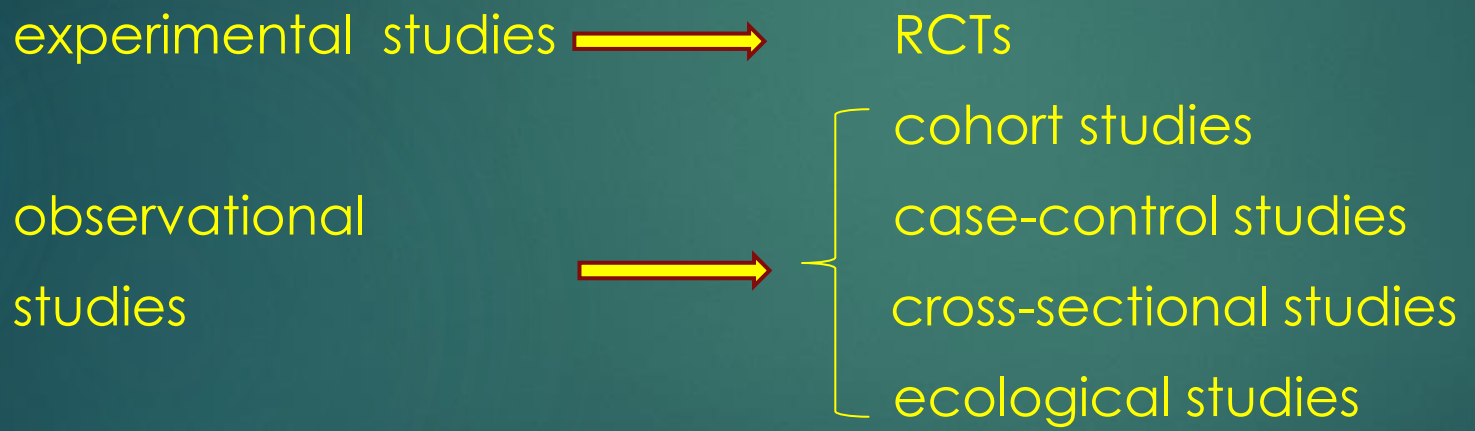
In our case, our data came from a randomised trial, i.e. the distribution of all the other characteristics of the participants are very similar (and each individual corresponded to 1000s of people), we could estimate the average treatment effect

$$ATE = 1/2 - 1/2 = 0$$

In real world data from observational studies, we need to account for confounders and many assumptions

Randomised Control Trials (RCTs)

RCTs are the gold standard in epidemiology and medical research to answer causal questions and understand what works for human health



Drugs would get approval in the market only if the RCT is successful (Phase III)

Problems with RCTs

- Usually strict inclusion criteria – might not be straightforward to extend inferences to other population groups
- Very expensive
- Impractical to conduct when timely decisions needed
- Unethical
- Relatively small with short follow-up time, hence difficult to measure "hard" outcomes (e.g. CVD or specific CVDs)

A promising alternative… use of observational data

The use of real-world data from observational studies is the only alternative solution in the absence of RCTs to assess causal inferences

ADVANTAGES : Using observational data, we have longer length of follow-up, and larger study size, sufficient power to focus on different population groups etc

PROBLEMS: How is/straightforward is that? What are the hurdles?

Observational studies for causal inference - assumption

- Observational studies (EHR, cohorts) can be used to draw causal inferences if certain assumptions are met. Those most invoked are:

- Consistency

- Conditional exchangeability (i.e. no unmeasured confounding)

- Positivity

Examples - Consistency

Consistency means that the observed outcome of a treated/untreated individual is the same as the counterfactual outcome, had he/she been treated or untreated. In other words $Y^{X=x} = Y$ for everyone with $X=x$

In other words, treatment should correspond to a well-defined intervention

 Bariatric surgery and mortality in people with obesity (well defined intervention)

 Weight change and mortality (how did people lose/gain weight? Physical activity? Diet?)

Examples - Conditional exchangeability

The conditional probability of receiving every value of treatment, depends only on measured covariates (not to unmeasured confounding)

 Bariatric surgery and mortality in people with obesity, when we have information on comorbidities, socioeconomic and lifestyle characteristics of the participants

 Use of statins and CVD, when we don't have information on the comorbidities of the participants

Examples - Positivity

The probability of receiving every value of treatment conditional on L is greater than zero (positivity) $0 < Pr(X \mid L = l) < 1$

 Bariatric surgery and mortality in people with obesity, when there are people from all the population groups (that define all potential strata) who underwent and didn't undergo bariatric surgery that some of them died and others remained alive

 Bariatric surgery and mortality in normal weight individuals, (nobody who is normal weight undergoes bariatric surgery)

Other assumptions (well known)

No interference: The exposure of one participant is not related with the outcome of another participant

Example of interference: An individual who, if not vaccinated, would have infected another person, but who, if vaccinated, would not infect that other person[VanderWeele et al 2015].

The infection (outcome) of the second individual depends on the treatment of the first individual

Other assumptions (well known)

- No Model misspecification
- Appropriately tackling missing data
- No selection bias

Other problems (under the radar)

Surprisingly, other sources of substantial bias (that are interconnected) are present when we want to estimate the causal effect of treatment using observation data

1. Specification of time zero
2. Include individuals based on post-baseline information for the definition of the exposure
3. Immortal time bias

Examples of “under the radar” biases

There are many “failures” from the analysis of observational data that have proven wrong by RCTs that where conducted afterwards

- Hormone therapy and CHD
- Antiretroviral therapy and HIV
- Statins and cancer

Why target trial emulation is so important

There are many "failures" from the analysis of observational data that have proven wrong by RCTs that where conducted afterwards

- Hormone therapy and CHD – much lower risk in postmenopausal women using therapy
- Antiretroviral therapy and HIV – risk of death doubled when deferring therapy just a few months
- Statins and cancer – 50% lower in people using statins

Why target trial emulation is so important

There are many "failures" from the analysis of observational data that have proven wrong by RCTs that where conducted afterwards

- Hormone therapy and CHD – much lower risk in postmenopausal women using therapy

[Grodstein F et al. NEJM 1996;335:453-461]

- Antiretroviral therapy and HIV – risk of death doubled when deferring therapy just a few months

[Kitahata M et al. NEJM 2009;360:1815-1826]

- Statins and colorectal cancer – 50% lower in people using statins

[Poynter J et al NEJM 2005; 352:2184-2192]

These are only examples of studies published in NEJM, the most influential medical journal!

Hormone therapy and CHD

- A number of studies that found that women taking hormone therapy had 60% less risk of developing CHD
- Prominent paper by NEJM 1996;335:453-461 → HR=0.68
- Guidelines were updated

American Heart Association (1996):
"Estrogen replacement therapy does look promising as a long-term protection against heart attack"

Hormone therapy and CHD

- A number of studies that found that women taking hormone therapy had 60% less risk of developing CHD
- Prominent paper by NEJM 1996;335:453-461 → HR=0.68
- Guidelines were updated

American Heart Association (1996):
"Estrogen replacement therapy does look promising as a long-term protection against heart attack"

- BUT… results from RCT (Manson et al. 2003) → HR=1.24
- Observational analysis using TTE framework → HR= 1.08 (Hernan 2008)

Antiretroviral therapy and HIV

- NEJM 2009;360:1815-1826 → – risk of death doubled when deferring therapy just a few months

- Observational analysis using TTE framework →

The benefits of immediate initiation of ART, such as prolonged survival and AIDS-free survival and increased virological suppression, were small in this high-income setting with relatively low CD4 count at HIV diagnosis (Lodi 2015)

Statins and colorectal cancer

- NEJM 2005;335:453-461 → 50% lower in people using statins

Statins and colorectal cancer

- NEJM 2005;335:453-461 → 50% lower in people using statins
- other observational studies also found lower in colorectal cancer in people using statins
- However, from more recent meta-analysis of RCTs, no effect was observed
- Observational analysis using TTE framework → no effect (Dickerman, 2019)

Statins and cancer

One observational study suggested a 77% lower risk of lung cancer among long-term statin users compared with nonusers (Khurana et al, 2007)

But…

- The investigators included individuals who were using statins before baseline (though they did not use pre-baseline therapy to quantify total duration of use)
- The investigators classified individuals based on their observed duration of statin therapy over follow-up (in this case, using postbaseline information to assign baseline treatment status).

Statins and cancer - Why TTE is helpful?

Interesting sensitivity analysis by (Dickerman 2019)

- When they utilised the TTE framework, they found no effect of statins on cancer
- When they implemented the (wrong) analysis of (Khurana et al, 2007), they found estimates identical with (Khurana et al, 2007), HR=0.23
- Observational data was not the problem, the problem was the data management and analysis

Statins and cancer

One observational study suggested a 77% lower risk of lung cancer among long-term statin users compared with nonusers (Khurana et al, 2007)

But…

- The investigators included individuals who were using statins before baseline (though they did not use pre-baseline therapy to quantify total duration of use)
- The investigators classified individuals based on their observed duration of statin therapy over follow-up (in this case, using postbaseline information to assign baseline treatment status).

Under the radar problems: Specification of time zero, include individuals based on post-baseline information for the definition of the exposure and immortal time bias

A way to address these biases… TTE

- All 3 examples we saw before were bias not because of the usual limitations of observational data (consistency, positivity, conditional exchangeability, measurement error, missing data, interference etc)
- conditional exchangeability, i.e. unmeasured confounding usually raises as the most important of them
- The problem is that they didn't utilise the framework of TTE and failed to provide evidence similar to a randomised trial

What is target trial emulation (TTE)?

TTE is a framework for comparative effectiveness research using observational data (cohorts, EHR etc) that makes the target trial explicit. Emulating the target trial involves the following steps (Hernan and Robins 2016, Hernan and Robins 2020)

1. We formulate the causal question
2. We explicitly specify the protocol of the target trial (eligibility criteria, treatment strategies, assignment procedures, follow-up period, outcome definition, causal contrast and analysis plan) we want to want to emulate, that is, the ideal trial we would like to conduct if it were possible. Useful to put this information in one table
3. We mimic each of the component of the target trial utilising our observational data. If this is not feasible, we need to reformulate the research question

Most common mistake: Specification of time zero (very important bias)

People usually compare individuals who are on a specific treatment vs individuals who are not on a specific treatment → WRONG

Most common mistake: Specification of time zero (very important bias)

People usually compare individuals who are on a specific treatment vs individuals who are not on a specific treatment → WRONG

We need to compare individuals who initiate a specific treatment vs individuals who do not initiate a specific treatment → CORRECT

References

- Dickerman, B. et al. Avoidable flaws in observational analyses: an application to statins and cancer. Nat Med 2019:25, 1601–1606
- Grodstein F et al. Postmenopausal estrogen and progestin use and the risk of cardiovascular disease. NEJM. 1996;335(7):453-61
- Hernán 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
- Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC
- Khurana, V et al. Statins reduce the risk of lung cancer in humans: a large case-control study of US veterans. Chest 2007:131, 1282–1288

References

- Kitahata M et al. Effect of Early versus Deferred Antiretroviral Therapy for HIV on Survival. NEJM 2009;360:1815-1826
- Lodi S, et al. Comparative effectiveness of immediate antiretroviral therapy versus CD4-based initiation in HIV-positive individuals in high-income countries: observational cohort study. Lancet HIV. 2015 Aug;2(8):e335-43
- Manson JE, et al. Estrogen plus progestin and the risk of coronary heart disease. N Engl J Med. 2003 Aug 7;349(6):523-34
- Poynter J et al. Statins and the Risk of Colorectal Cancer NEJM 2005; 352:2184-2192
- VanderWeele TJ, et al. Interference and Sensitivity Analysis. Stat Sci. 2014;29(4):687-706

Thanks!

E-mail: m.katsoulis@ucl.ac.uk

Twitter: @MK_statistics