The Language of Causal Inference

Seoyoung C. Kim, MD, ScD, MSCE
Director, PRIME
Associate Professor of Medicine
Division of Pharmacoepidemiology and Pharmacoeconomics
Division of Rheumatology, Inflammation, and Immunity
Brigham and Women's Hospital and Harvard Medical School

Disclosures

- NIH- NIAMS:
  - R01- AR069557, R01-AR073314, P30-AR072577
- Research grants to the Brigham and Women's Hospital
  - Bristol-Myers Squibb, Genentech/Roche, AbbVie, and Pfizer

Learning objectives

1. Define the role of causal inference within observational studies
2. Identify the common types of studies and methods used to examine causal inference
3. Define common vocabulary associated with causal inference
4. Outline the strengths and limitations of the methodologies for causal inference

References

- Causal Inference: What If (Hernan and Robins): https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/
- https://doi.org/10.1093/ije/dyt127
- https://www.nature.com/articles/nrrheum.2015.30
- https://academic.oup.com/aje/article/158/9/915/102549
- https://onlinelibrary.wiley.com/doi/abs/10.1002/pds.1357
- https://pubmed.ncbi.nlm.nih.gov/33385861/
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3731075/

Causal inference

- Identification and estimation of causal effects in populations
  - *i.e.*, numerical quantities that measure changes in the distribution of an outcome under different interventions

- Individual causal effects
  - For specific individual 'i':

Causal effect for individual i:
$Y_i^{a=1} \neq Y_i^{a=0}$

- Average causal effects
  - In the the population :

$$\mathrm{E}[Y^{a=1}] \neq \mathrm{E}[Y^{a=0}]$$

Measure of causal effects

- Causal risk difference (RD)
- Causal risk ratio (RR)
- Causal odds ratio (OR)

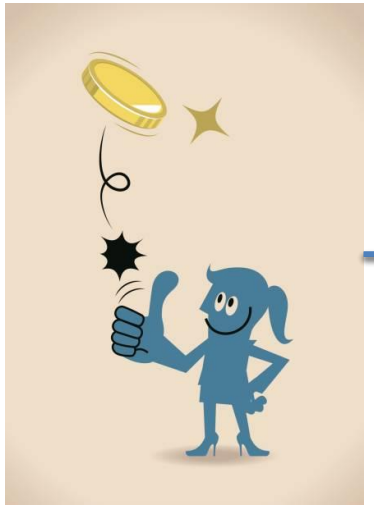
(i) $\Pr[Y^{a=1}=1] - \Pr[Y^{a=0}=1] = 0$

(ii) $\dfrac{\Pr[Y^{a=1}=1]}{\Pr[Y^{a=0}=1]} = 1$

(iii) $\dfrac{\Pr[Y^{a=1}=1]/\Pr[Y^{a=1}=0]}{\Pr[Y^{a=0}=1]/\Pr[Y^{a=0}=0]} = 1$

Let’s conduct an RCT of RA treatment A

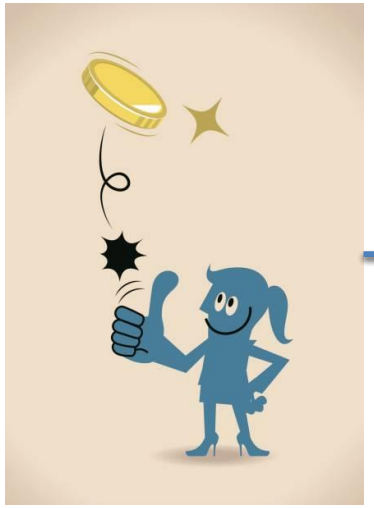
Randomization



50:50

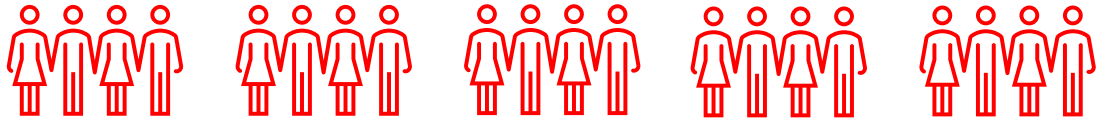
RCT: Treatment assignment

Randomization



50:50

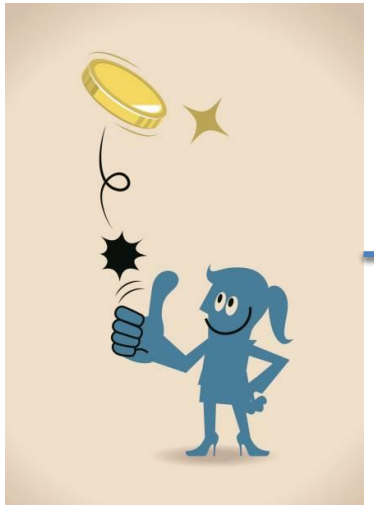
Treatment A



No treatment

RA remissions occurred (Y=1) in RCT

Randomization



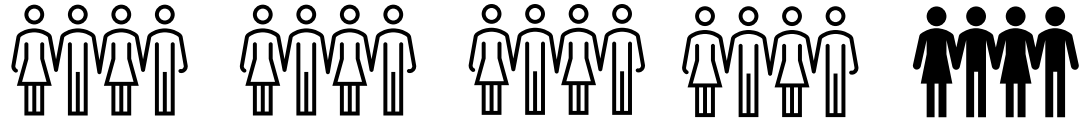
50:50

Treatment A: A=1

Probability of remission: 0.4

$Pr[Y^{A=1}=1] = 0.4$

No treatment: A=0



Probability of remission: 0.2

$Pr[Y^{A=0}=1] = 0.2$

RD: 0.2

RR: 2.0

How do we interpret RD=0.2 or RR=2.0 in this RCT?

- To infer a causal effect of exposure on outcome, we need to know the risk of the outcome in the presence of exposure AND in the absence of exposure given all other conditions being same

Counterfactual outcome

How do we interpret RR=2.0 in that RCT?

- To infer a causal effect of exposure on outcome, we need to know the risk of the outcome in the presence of exposure AND in the absence of exposure given all other conditions being same

Randomization➔ Exchangeability between the 2 groups

How do we interpret RD=0.2 or RR=2.0 in this RCT?

- To infer a causal effect of exposure on outcome, we need to know the risk of the outcome in the presence of exposure AND in the absence of exposure given all other conditions being same

- RD= 0.2 or RR=2.0 in the RCT of Treatment A: is the causal effect of the treatment on RA remission

  *Number needed to treat (NNT): the average number of individuals that need to receive treatment (A=1) to achieve the number of outcomes (Y=1) by 1

  $NNT = 1/RD = 1/ \{Pr[Y^{A=1}=1] - Pr[Y^{A=0}=1]\}$

  RD=0.2, NNT=5: on average, one needs to treat 5 RA patients with A to achieve 1 remission

But we don’t always have ‘randomization’

Causal inference in non-randomized (observational) studies

- Exposure occurs temporally prior to outcomes → cohort or nested case-control studies ~~(No cross-sectional or case series)~~
- Individuals are either exposed or unexposed but not both
  - No counterfactual outcome available (unless we have a time machine or can clone individuals)
- No randomization → Non-exchangeability
- Then how can we make causal inference about the exposure on outcome in non-randomized settings?

Key concepts in designing observational studies for causal inference

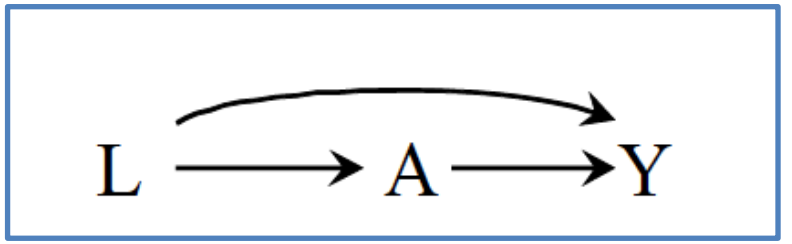
- Ideally, we would like to compare the exposed to the same exact people had they not been exposed, but we don't have a time machine or cloning technology
- So we need to use the unexposed group that is presumed to be 'identical' to the exposed group with respect to their risk of outcomes
  - "Conditional exchangeability" –the exposed and the unexposed are exchangeable within strata of the measured covariates (within levels of L), also assuming no unmeasured confounding
  - "Consistency" – if A_i=a, then $Y_i^a = Y_i^A = Y_i$: an individual treated with 'a' has an outcome Y equal to his counterfactual outcome Y^a
  - "Positivity" – the conditional probability of receiving every value of treatment is non-zero (>0): *e.g., violation of positivity: drug contraindication, no appropriate match, small sample sizes and/or high-dimensional data*

Key concepts in designing observational studies for causal inference

- Well-defined interventions: One must know which treatments are being compared
- For each causal effect of interest, we can imagine a hypothetical RCT to test that hypothesis
  - More straightforward if you can design a 'target trial' (more in the next session)
  - When conducting the target trial is not feasible, or ethical, or timely, we can consider causal analyses of observational data to emulate the target trial

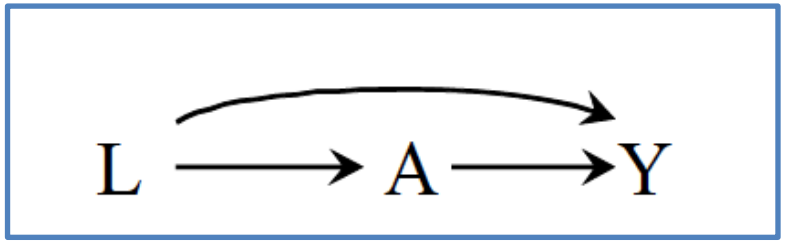
Graphical representation of causal effects

- We can conceptualize problems using causal diagrams like directed acyclic graphs (DAGs)



- Causal DAGs are a way to be explicit about your assumptions about the causal structure of the relationships between exposures, potential confounders and their consequences.

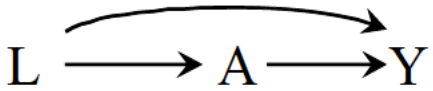
Causal DAGs



- Directed: time flows from left to right via directed arrows
  - L is temporally prior to A and Y, and A is temporally prior to Y
  - *L is a common cause of A and Y ➔ L: confounder (lack of exchangeability)*
- Acyclic: no cycles/feedback loops, a variable cannot cause itself, either directly or through another variable
- Causal Graphs:
  - An arrow from one node to another indicates 'a direct causal effect'

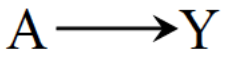
Examples of causal DAGs

- Conditionally randomized experiment:



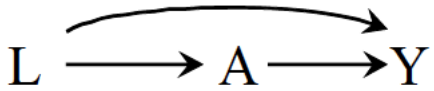
  - e.g., Patients are randomly assigned to Treatment A depending on the RA disease activity (L)

- Marginally randomized experiment:



  - e.g., Patients are randomly assigned to Treatment A regardless of their RA disease activity

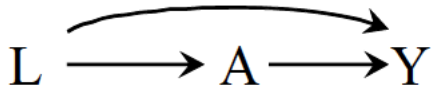
Common cause: confounding



- Observational study (same DAG as conditional RCT):
  - e.g., Patients are prescribed Treatment A depending on the RA disease activity (L)
  - Assuming the treatment A depends solely on L, and no other cause of Y exists
  - 2 sources of associations:
    - Path A → Y : causal effect of A on Y
    - Path A → L → Y : backdoor path through their common cause L
      - In the presence of L, association is NOT causation
      - Must account for L in an analysis to obtain a valid (unbiased) estimate of the A-Y relationship

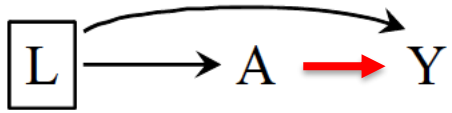
Conditioning on a common cause

- Observational study:



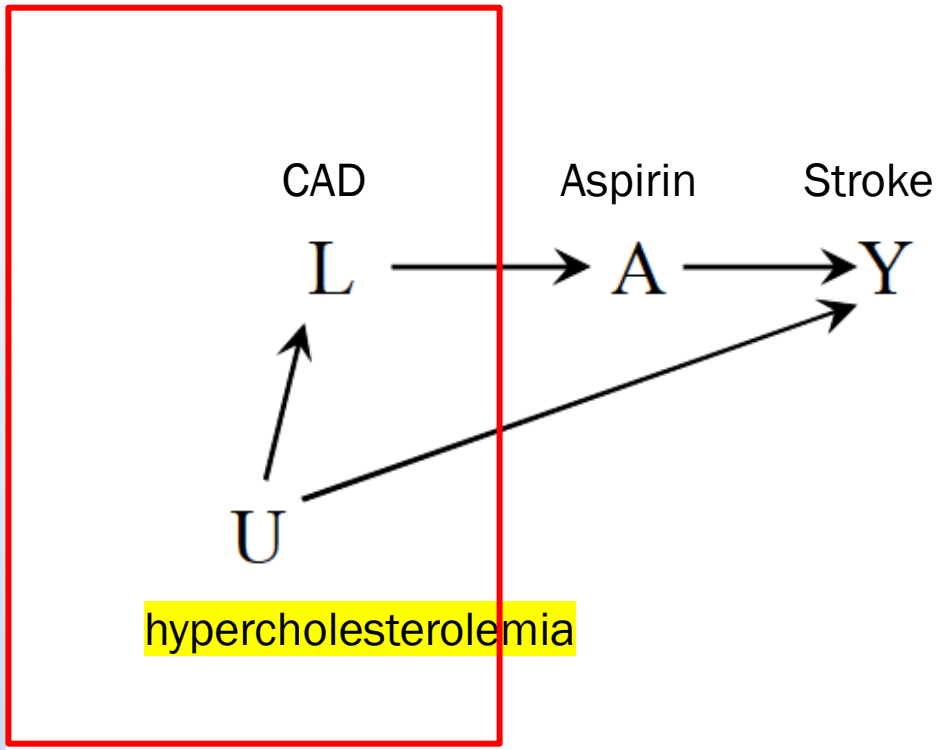
  - e.g., Patients are prescribed Treatment A depending on the RA disease activity (L)
  - Assuming the treatment A depends solely on L, and no other cause of Y exists

- Conditioning on a common cause L:

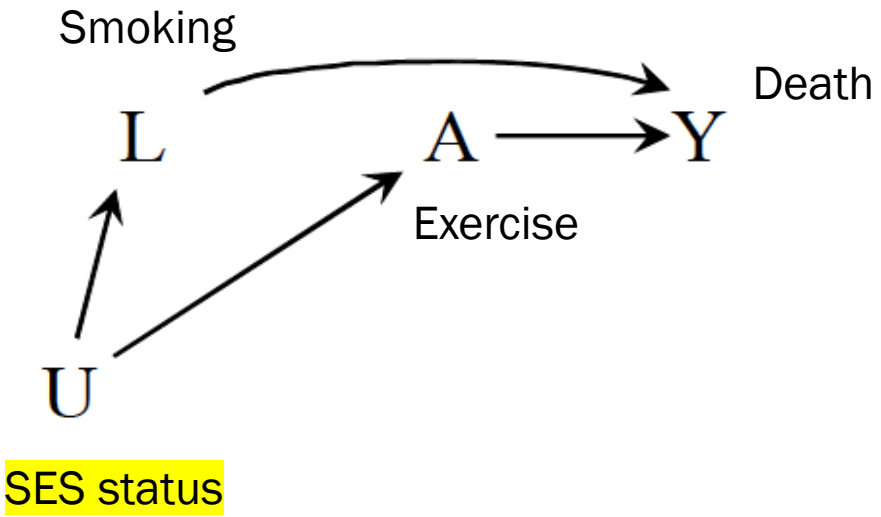


  - Implies conditional exchangeability within the level of L: blocks the backdoor path
  - Assuming no unmeasured confounding, only 1 path from A to Y exists: causal

Example of unmeasured confounding (U)

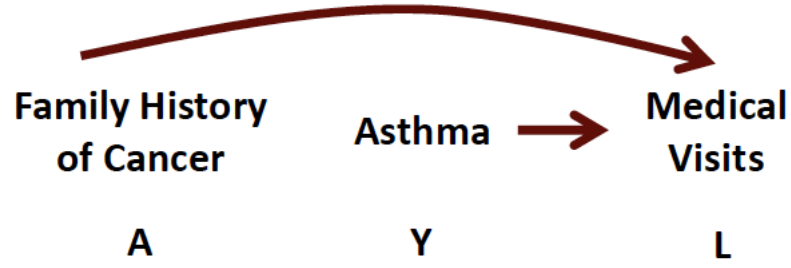


Confounding by indication

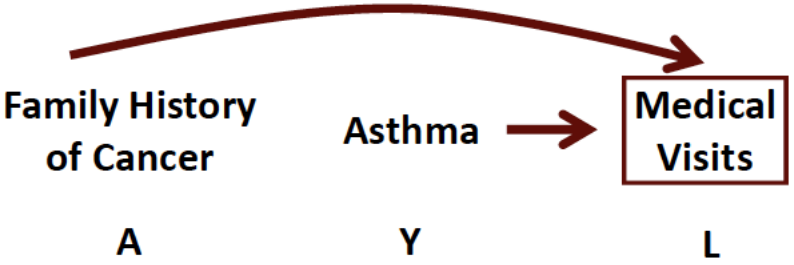


Conditioning on a common effect

- Common effect: L is a common effect of and Y
  - L is a collider on the path A → L ← Y
  - No association between A and Y because the collider L blocks the flow of association



- Conditioning on a common effect (collider) L
  - Backdoor path is open from Y → L → A
  - Leading to selection bias



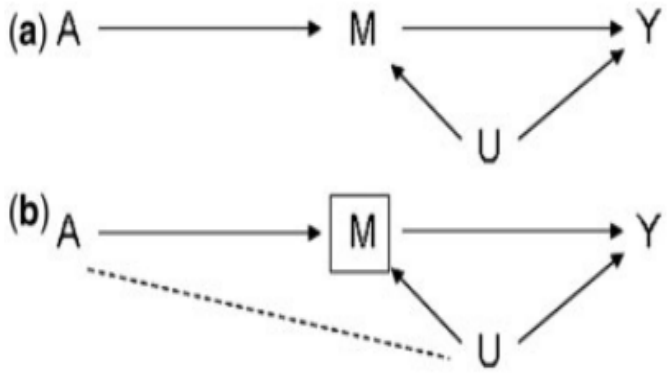
Conditioning on a common cause vs. a common effect

- Conditioning on a common cause removes confounding.
- Conditioning on a common effect introduces selection bias.
- Therefore, we need expert knowledge to determine if we should adjust for a variable.
  - We need to set the timeline of all these variable measurement correctly.
  - Statistical criteria are insufficient to characterize confounding and confounders.

Mediator

Mediator

- What happens if one controls for the mediator M?
- In the presence of mediator-outcome confounding, conditioning on M induces an association through a backdoor path: Y → U→ M→ A



Causal DAGs

- We use DAGs to conceptualize problems and help us be explicit about the underlying assumptions.
- Your proposed DAG is generally simplified and may or may not correctly represent the true state of nature.
- The analysis that is dictated by a particular DAG estimates the causal effect of exposure only if the DAG is correct.

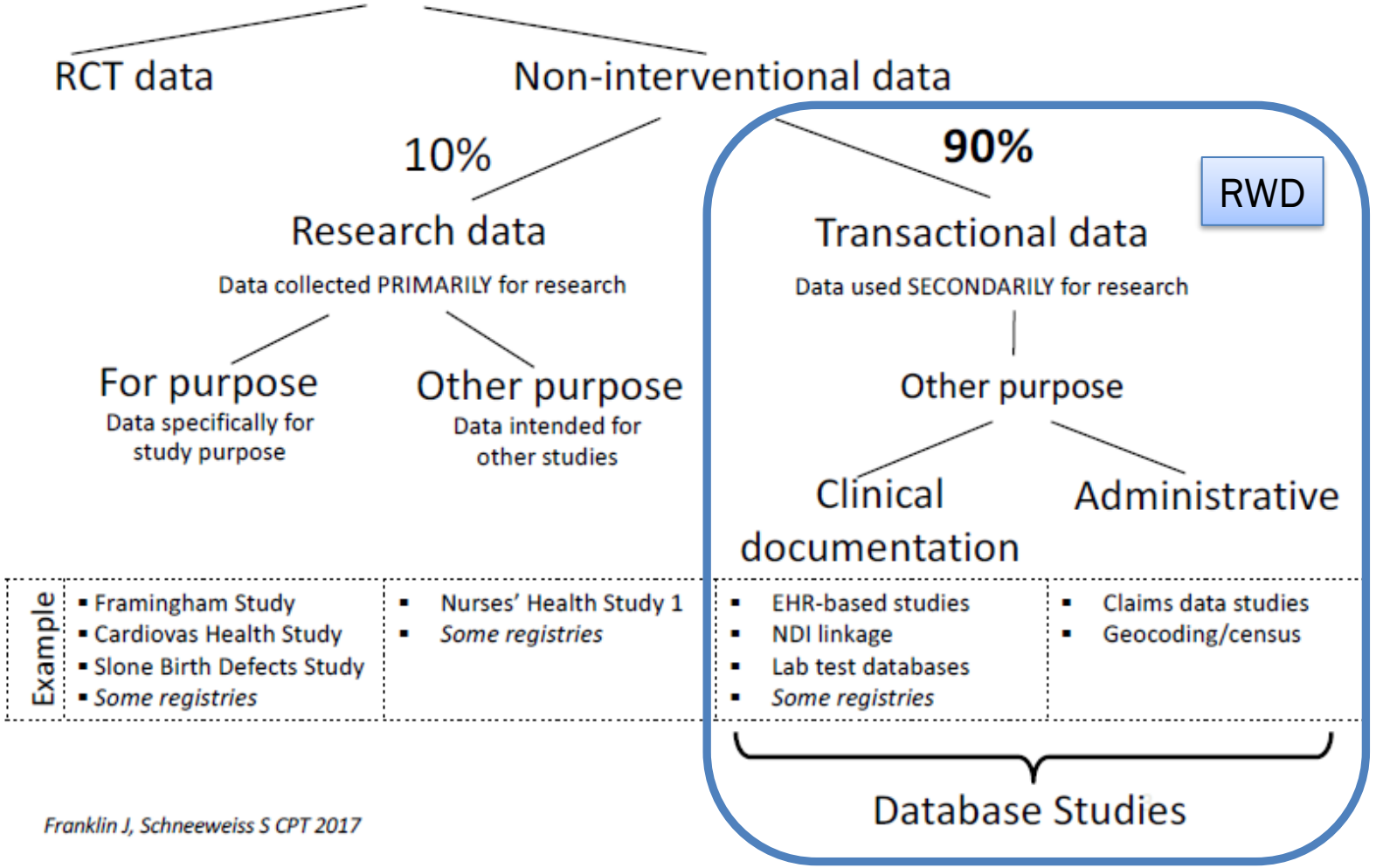
A hypothetical study question

- Let's consider conducting an observational cohort study to assess the effect of methotrexate (MTX) on the risk of myocardial infarction (MI) among RA patients
  - % MI among RA patients exposed MTX is observable
  - % MI among RA patients not exposed to MTX is observable
  - But are they comparable/exchangeable?
    - In a non-randomized setting, 'RA patients treated with MTX' may have different MI risk factors versus 'RA patients not exposed to MTX'.
      - LACK of EXCHANGEABILTY or CONFOUDNING
    - Then % MI among RA patients not exposed to MTX is not a good proxy for the conditional counterfactual outcome: you cannot infer causal effect validly

Study question

- Does MTX decrease the risk of MI among RA patients ?
  - Study population: RA patients
  - Data sources: real-world data sources

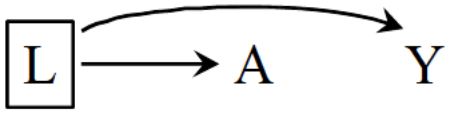
Effectiveness Research with Healthcare Databases



Franklin J, Schneeweiss S CPT 2017

Study question

- Does MTX decrease the risk of MI among RA patients ?
  - Study population: RA patients
  - Data sources: real-world data sources - registries, EMR, claims
  - Exposure of interest:
    - 'Methotrexate' versus 'No methotrexate'
    - Are these two groups comparable? We must control for confounding to ensure 'conditional exchangeability'



How to control for confounding ?

- By study design
  - New user design
  - Active comparator
- By statistical methods
  - Multivariable adjustment
  - Stratification
  - Standardization
  - Propensity score methods

Observational data on MTX and risk of CVD

Systematic Review and Meta-Analysis of *Methotrexate* Use and Risk of Cardiovascular Disease

Renata Micha, RD, PhD[a,*], Fumiaki Imamura, PhD[a], Moritz Wyler von Ballmoos, MD, PhD[b], Daniel H. Solomon, MD[c], Miguel A. Hernán, MD, DrPH[a,e], Paul M. Ridker, MD[d], and Dariush Mozaffarian, MD, DrPH[a,d]

Am J Cardiol 2011

Limited by non-user comparator +/- prevalent user design

Table 1
Identified studies evaluating methotrexate use in patients with systemic inflammation and occurrence of cardiovascular disease‡

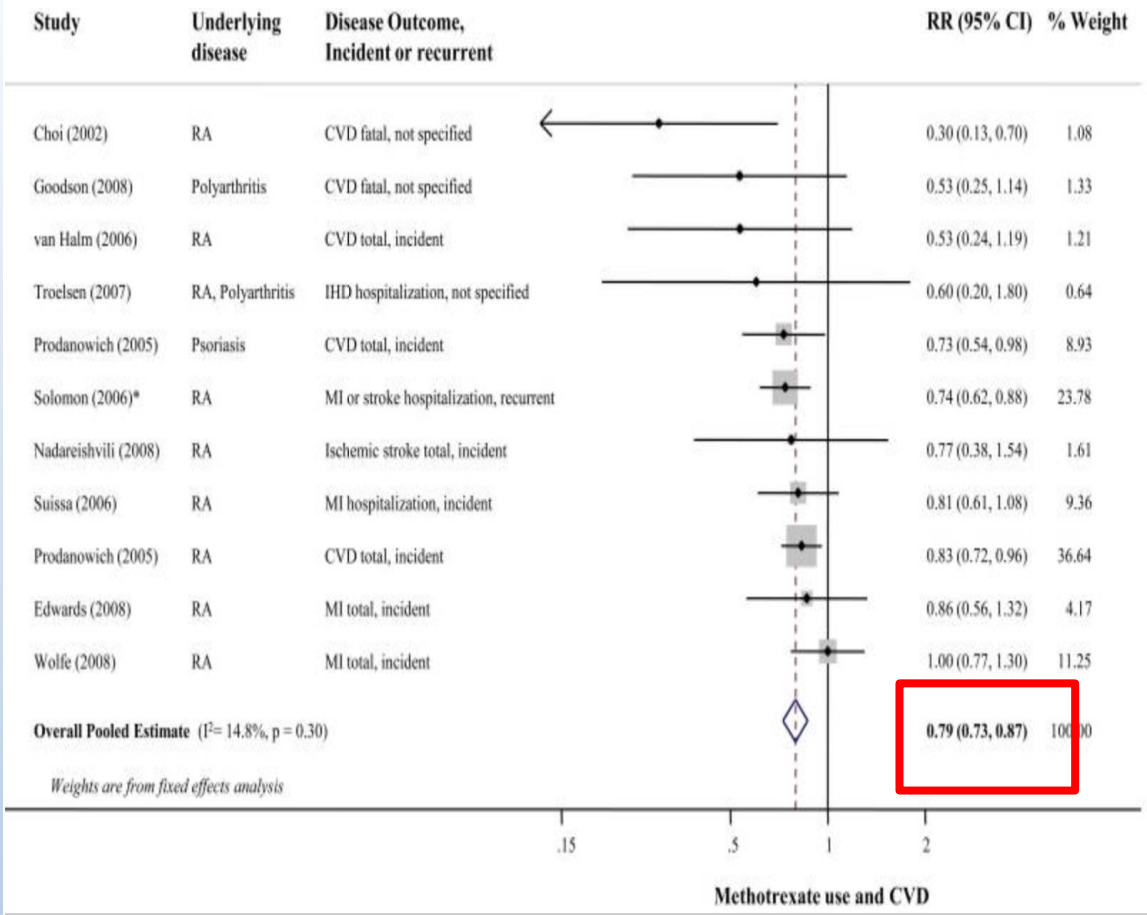
Study	Country	Study Name	Underlying Disease	Ascertainment of MTX Use	MTX Comparison Groups
Prospective cohort studies					
Choi et al (2002)[14]	United States	Wichita Arthritis Center Cohort	RA	Arthritis medical record database	Initiators vs noninitiators
Solomon et al (2006)[10]	United States	Pharmaceutical Assistance Contract for the Elderly	RA	Health care utilization database	Initiators vs noninitiators
Suissa et al (2006)[15]	North America	PharMetrics Patient-Centric Outcomes Database Cohort	RA	Health care utilization database	Current users vs noncurrent users
Troelsen et al (2007)[16]	Denmark	NA	RA	Clinical chart	Current users vs noncurrent users
Nadareishvili et al (2008)[17]	United States	National Data Bank for Rheumatic Disease Longitudinal Study	RA	Patients' self-report	Ever users vs never users
Wolfe et al (2008)[18]	United States	National Data Bank for Rheumatic Disease Longitudinal Study	RA	Patients' self-report	Ever users vs never users
Edwards et al (2008)[20]	United Kingdom	United Kingdom General Practice Research Database	RA	General practice database	Ever users vs never users
Goodson et al (2008)[21]	United Kingdom	United Kingdom Norfolk Arthritis Register	Polyarthritis	Not reported	Current users vs never users
Retrospective cohort studies					
Prodanowich et al (2005)[19]	United States	Miami Veterans Cohort	Psoriasis RA	Pharmacy database	Ever users vs never users
van Halm et al (2006)[13]	The Netherlands	Jan van Breemen Institute	RA	Medical record	Ever users vs never users

Observational data on MTX and risk of CVD

Systematic Review and Meta-Analysis of *Methotrexate* Use and Risk of Cardiovascular Disease

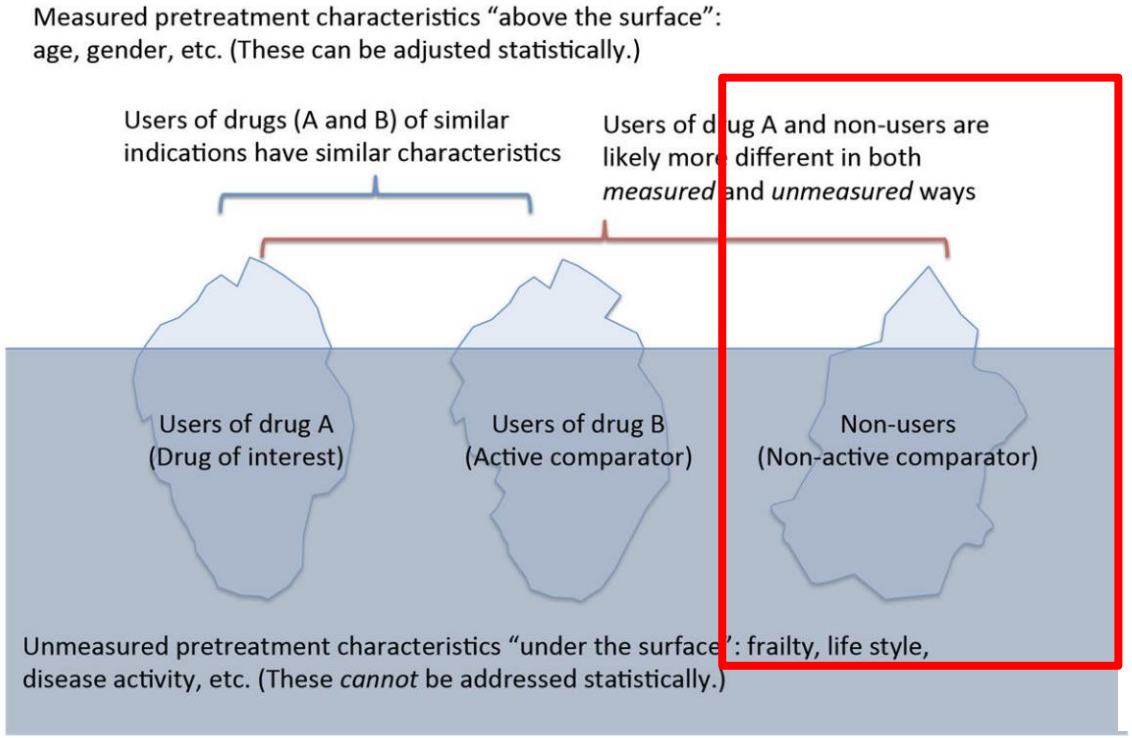
Renata Micha, RD, PhD[a,*], Fumiaki Imamura, PhD[a], Moritz Wyler von Ballmoos, MD, PhD[b], Daniel H. Solomon, MD[c], Miguel A. Hernán, MD, DrPH[a,e], Paul M. Ridker, MD[d], and Dariush Mozaffarian, MD, DrPH[a,d]

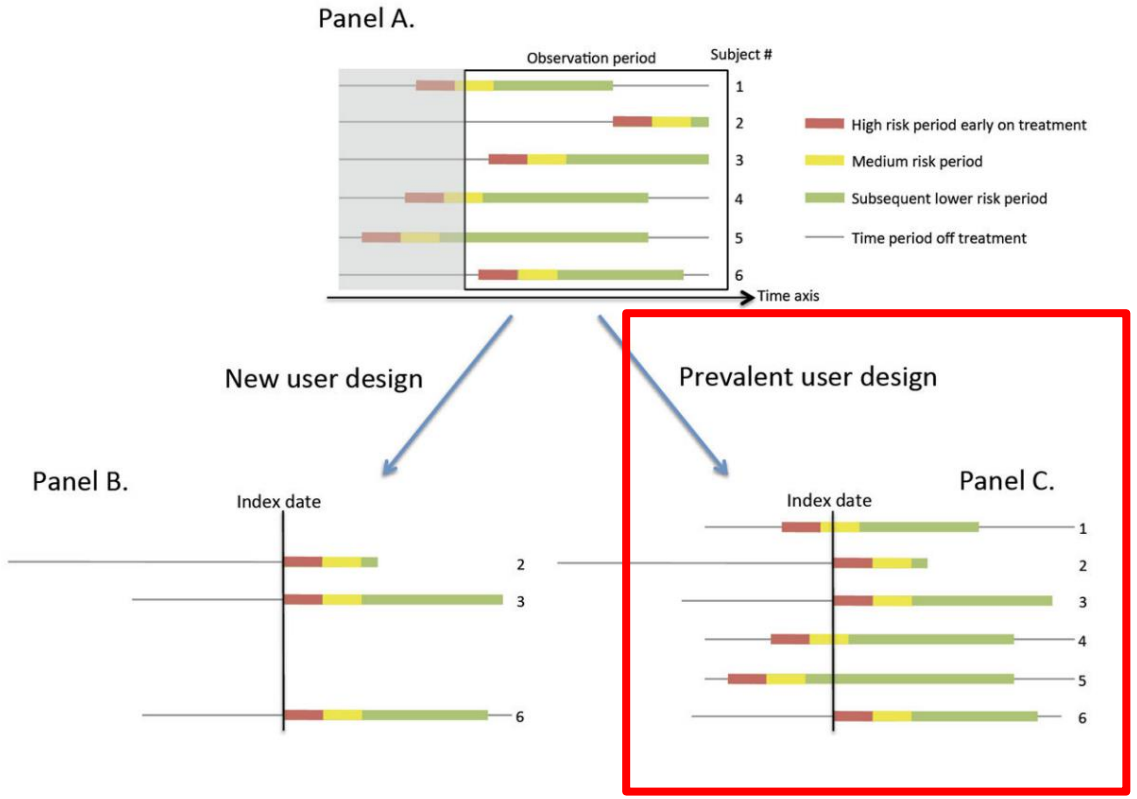
Am J Cardiol 2011



In summary, our findings provide support for the inflammatory hypothesis of atherothrombosis. Given the heterogenous methods identified, future observational studies should ideally compare methotrexate initiators versus noninitiators and also adjust for confounding by underlying disease severity. Our findings also support the need for further experimental studies to elucidate likely mechanisms and randomized clinical trials to establish causality.

Why non-user comparator/prevalent user design is bad?





Yoshida (2015) DOI: 10.1038/nrrheum.2015.30

Well-known example of prevalent user/non-user bias

The New England Journal of Medicine

VOLUME 335 AUGUST 15, 1996 NUMBER 7

POSTMENOPAUSAL ESTROGEN AND PROGESTIN USE AND THE RISK OF CARDIOVASCULAR DISEASE

FRANCINE GRODSTEIN, SC.D., MEIR J. STAMPFER, M.D., JOANN E. MANSON, M.D., GRAHAM A. COLDITZ, M.B., B.S., WALTER C. WILLETT, M.D., BERNARD ROSNER, PH.D., FRANK E. SPEIZER, M.D., AND CHARLES H. HENNEKENS, M.D.

TABLE 2. RELATIVE RISK OF CARDIOVASCULAR DISEASE AMONG CURRENT USERS OF CONJUGATED ESTROGEN ALONE OR WITH PROGESTIN AS COMPARED WITH NONUSERS, 1978 TO 1992.*

HORMONE USE	PERSON-YEARS	MAJOR CORONARY DISEASE			STROKE (ALL TYPES)		
		NO. OF CASES	RELATIVE RISK (95% CI)		NO. OF CASES	RELATIVE RISK (95% CI)	
			Age Adjusted	*Multivariate Adjusted†*		*Age Adjusted*	*Multivariate Adjusted†*
Never used	304,744	431	1.0		270	1.0	
Currently used							
Estrogen alone	82,626	47	0.45 (0.34–0.60)	0.60 (0.43–0.83)	74	1.13 (0.88–1.46)	1.27 (0.95–1.69)
Estrogen with progestin	27,161	8	0.22 (0.12–0.41)	0.39 (0.19–0.78)	17	0.74 (0.45–1.20)	1.09 (0.66–1.80)

*CI denotes confidence interval.

†The analysis was adjusted for age (in five-year categories), time (in two-year categories), age at menopause (in two-year categories), body-mass index (in quintiles), diabetes (yes or no), high blood pressure (yes or no), high cholesterol level (yes or no), cigarette smoking (never, formerly, or currently [1 to 14, 15 to 24, or 25 or more cigarettes per day]), past oral-contraceptive use (yes or no), parental history of myocardial infarction before the age of 60 years (yes or no), and type of menopause (natural or surgical).

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 AUGUST 7, 2003 VOL. 349 NO. 6

Estrogen plus Progestin and the Risk of Coronary Heart Disease

JoAnn E. Manson, M.D., Dr.P.H., Judith Hsia, M.D., Karen C. Johnson, M.D., M.P.H., Jacques E. Rossouw, M.D., Annlouise R. Assaf, Ph.D., Norman L. Lasser, M.D., Ph.D., Maurizio Trevisan, M.D., Henry R. Black, M.D., Susan R. Heckbert, M.D., Ph.D., Robert Detrano, M.D., Ph.D., Ora L. Strickland, Ph.D., Nathan D. Wong, Ph.D., John R. Crouse, M.D., Evan Stein, M.D., and Mary Cushman, M.D., for the Women's Health Initiative Investigators*

Table 1. Coronary Outcomes among Women Randomly Assigned to Estrogen plus Progestin, as Compared with Those Assigned to Placebo.*

Variable	Estrogen-plus-Progestin Group (N=8506)	Placebo Group (N=8102)	Adjusted Hazard Ratio	Nominal 95% CI	Adjusted 95% CI
Mean follow-up time (mo)	67.8	66.8			
	no. of cases (annualized percentage)				
CHD	188 (0.39)	147 (0.33)	1.24	1.00–1.54	0.97–1.60
Nonfatal MI					
Including silent MI	151 (0.31)	114 (0.25)	1.28	1.00–1.63	0.96–1.70
Excluding silent MI	147 (0.31)	109 (0.24)	1.30	1.01–1.67	0.97–1.74
Death due to CHD	39 (0.08)	34 (0.08)	1.10	0.70–1.75	0.65–1.89

(*Epidemiology* 2008;19: 766–779)

ORIGINAL ARTICLE

Observational Studies Analyzed Like Randomized Experiments

An Application to Postmenopausal Hormone Therapy and Coronary Heart Disease

Miguel A. Hernán,[a,b] *Alvaro Alonso,*[c] *Roger Logan,*[a] *Francine Grodstein,*[a,d] *Karin B. Michels,*[a,d,e] *Walter C. Willett,*[a,d,f] *JoAnn E. Manson,*[a,d,g] *and James M. Robins*[a,h]

TABLE 3. Estimates of the Intention-to-Treat Effect of Initiation of Estrogen/Progestin Therapy on the Incidence of CHD Events in the NHS "Trials"

		Follow-Up Period	
	All	**0–24 Mo**	**>24 Mo**
Initiators			
Total no.	7258	7258	7221
No. CHD events	98	22	76
Noninitiators			
Total no.	141,002	141,002	139,599
No. CHD events	3606	512	3094
	HR (95% CI)	**HR (95% CI)**	**HR (95% CI)**
All women	0.96 (0.78–1.18)	1.42 (0.92–2.20)	0.88 (0.69–1.12)
By time after menopause (y)			
<10	0.84 (0.61–1.14)	1.33 (0.66–2.64)	0.77 (0.54–1.09)
≥10	1.12 (0.84–1.48)	1.48 (0.83–2.64)	1.05 (0.77–1.43)
P for interaction	0.08	0.90	0.07
By age (y)			
<60	0.86 (0.65–1.14)	1.36 (0.73–2.52)	0.78 (0.57–1.07)
≥60	1.15 (0.85–1.57)	1.49 (0.79–2.80)	1.08 (0.76–1.54)
P for interaction	0.05	0.72	0.06

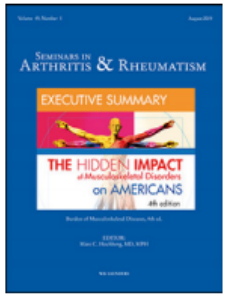
Immortal time bias

Seminars in Arthritis and Rheumatism 51 (2021) 211–218

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit



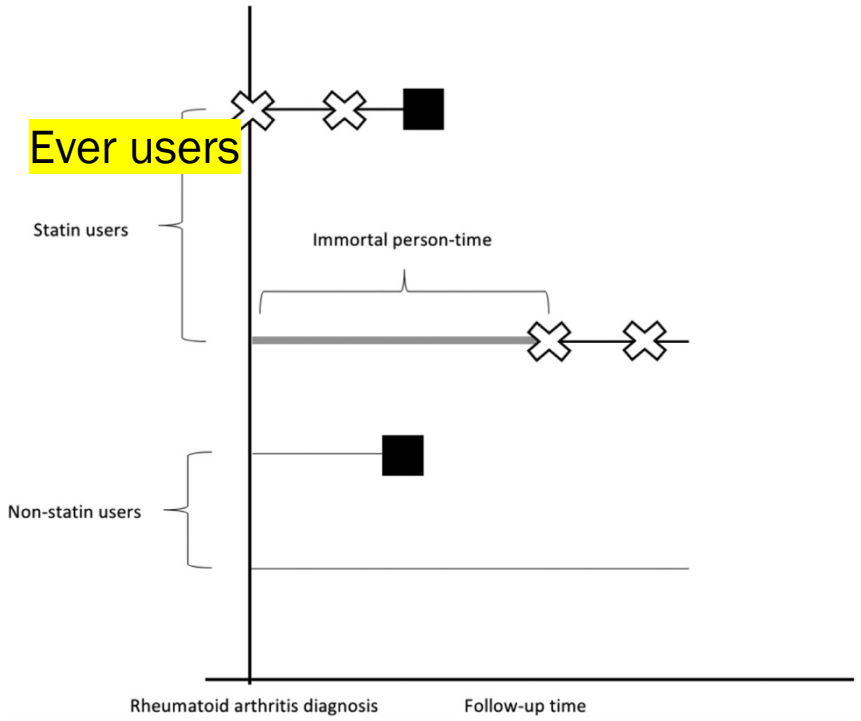
Statins and lower mortality in rheumatic diseases: An effect of immortal time bias?

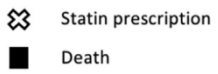
Devin Abrahami[a,b], Marie Hudson[a,c,d], Samy Suissa[a,b,d,*]

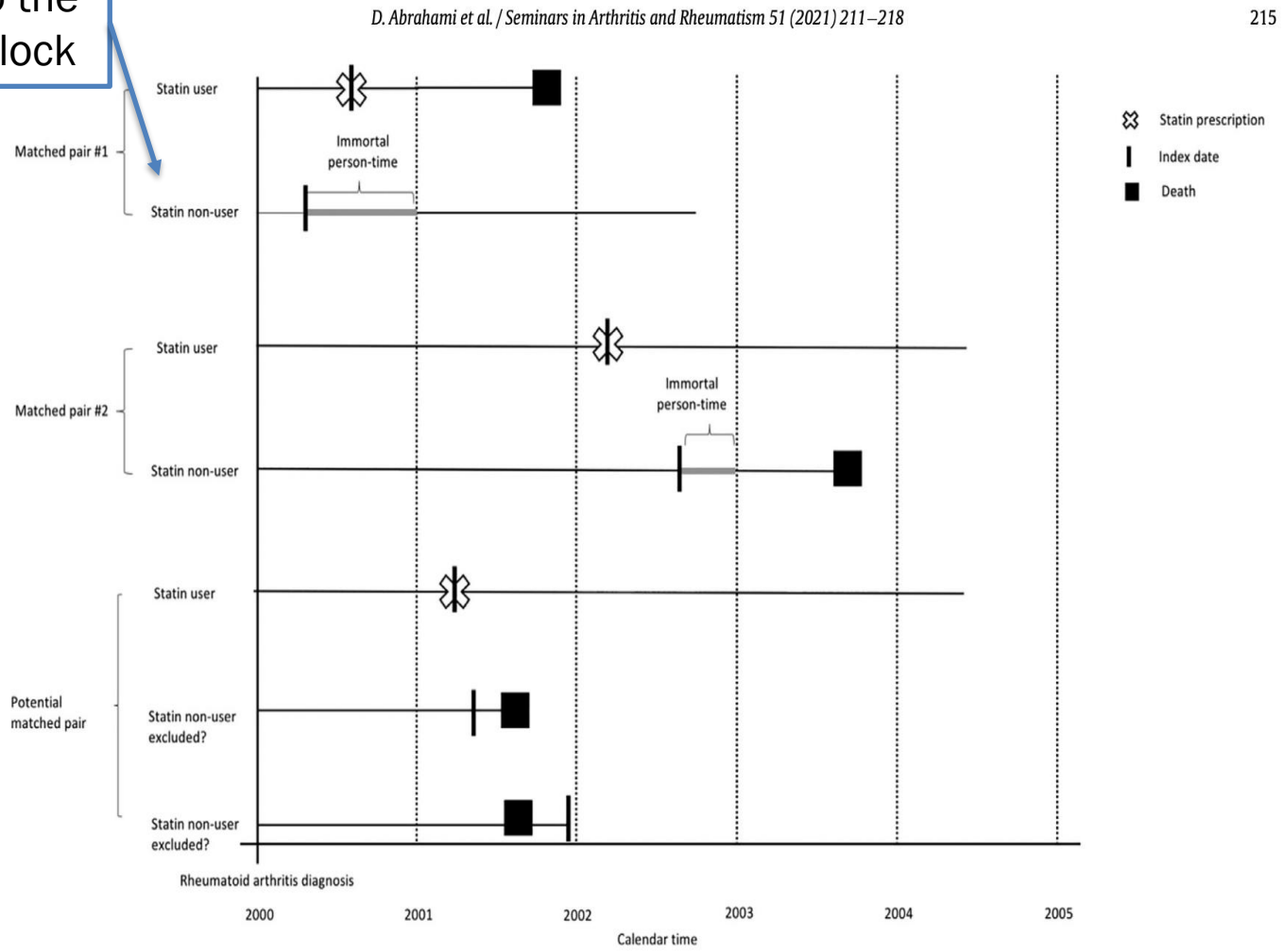
[a] Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, Canada
[b] Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada
[c] Division of Rheumatology, Jewish General Hospital, Montreal, Canada
[d] Department of Medicine, McGill University, Montréal, Canada

Types of immortal time bias

Non-users had to survive to the end of the cohort accrual block







Classical immortal time bias

Random-calendar date immortal time bias

Study question

- Does MTX decrease the risk of MI among RA patients ?
  - Study population: RA patients
  - Data sources: real-world data sources - registries, EMR, claims
  - Exposure of interest:
    - ‘Methotrexate’ versus ‘No methotrexate’
  - Outcome of interest: MI
  - Statistical analysis:
    - Intention-to-treat analysis
    - Per-protocol analysis

Summary

- When we cannot conduct an RCT for whatever reasons, well-designed/executed observational analyses can be useful for causal inference
- Estimation of causal effects in observational studies is challenging but doable with careful methodological considerations
- Three important assumptions to remember for causal inference
  - Conditional exchangeability, Consistency, and Positivity
- Defining treatments clearly (imagine a hypothetical target trial)
  - Will help determine 'time zero' and set up the timeline of the study correctly
  - Commonly used 'current users, never or ever users' are almost never used in RCTs. [DO NOT USE IT in your study]

Thank you!

Email: sykim@bwh.harvard.edu