USING BIG DATA TO EMULATE A TARGET TRIAL WHEN A RANDOMIZED TRIAL IS NOT AVAILABLE

Article written by Hernan MA and Robins JM in American Journal of Epidemiology 2016: 183 (8); 758-764

Presented by Dr. Chusak Limotai

“Target experiment” or “Target trial”

[Target trial (RCT): RCT which we target, but use observational data]

- Ideally, questions about comparative effectiveness or safety would be answered using an appropriately designed and conducted randomized experiment

- When we cannot conduct a randomized experiment,
  - *We can use observational analyses of existing data because the randomized trial that would answer our causal question—the target trial—is not feasible, ethical, and timely*
  - ***We aim to perform causal inference from large observational databases (big data) in an attempt to emulate a randomized experiment -“Target experiment” or “Target trial” –***

Ideal objective of analysis of big data

- If the emulation is successful, the analysis of the observational data yields the same effect estimates (**except for random variability**) as the target trial would have yielded had the latter been conducted

(**Emulate = imitate; copy**)

Big data

- Large observational databases are often used to answer questions about comparative effectiveness or safety
- Typically include many variables measured in many people
- The increasing availability of big data facilitates the emulation of target trials

Authors' aims

- Though the concept of a target trial is implicit in many big data analyses, the target trial itself is rarely characterized
- Authors outline a framework for comparative effectiveness research using big data that revolves around the explicit description and emulation of the target trial
- This framework channels the existing **"counterfactual theory"** for
  - *comparing the effects of point treatments and sustained treatment strategies*
  - *organizes analytic approaches dispersed throughout the literature*
  - *provides a structured process for the criticism of observational studies*
  - *helps avoid common methodologic pitfalls*

Counterfactual theory in causal reasoning

Counterfactual (adj.) = thinking about what did not happen but could have happened, or relating to this kind of thinking [Cambridge Dictionary]

- In a trial with two arms, for example, randomization causes the two treatment arms to be good substitutes (in probability) for the **two counterfactual (assumptions) disease frequencies**:
  - *What would have happened if everyone enrolled in the trial had received the treatment, and*
  - *What would have happened if everyone enrolled in the trial had received the placebo*
- A comparison of disease frequencies across treatment groups tells us (within sampling error) if (and by how much) disease frequency in the target population would have differed under the different treatments

Maldonado G; Annals of Epidemiology 2016: 26; 681-682

Scenario in this review

- Suppose we want to estimate the effect of estrogen plus progestin hormone therapy on the 5-year risk of breast cancer among postmenopausal women.

- Using a large database of health care claims

The protocol of a target trial to estimate the effect of postmenopausal hormone therapy on the 5-year risk of breast cancer

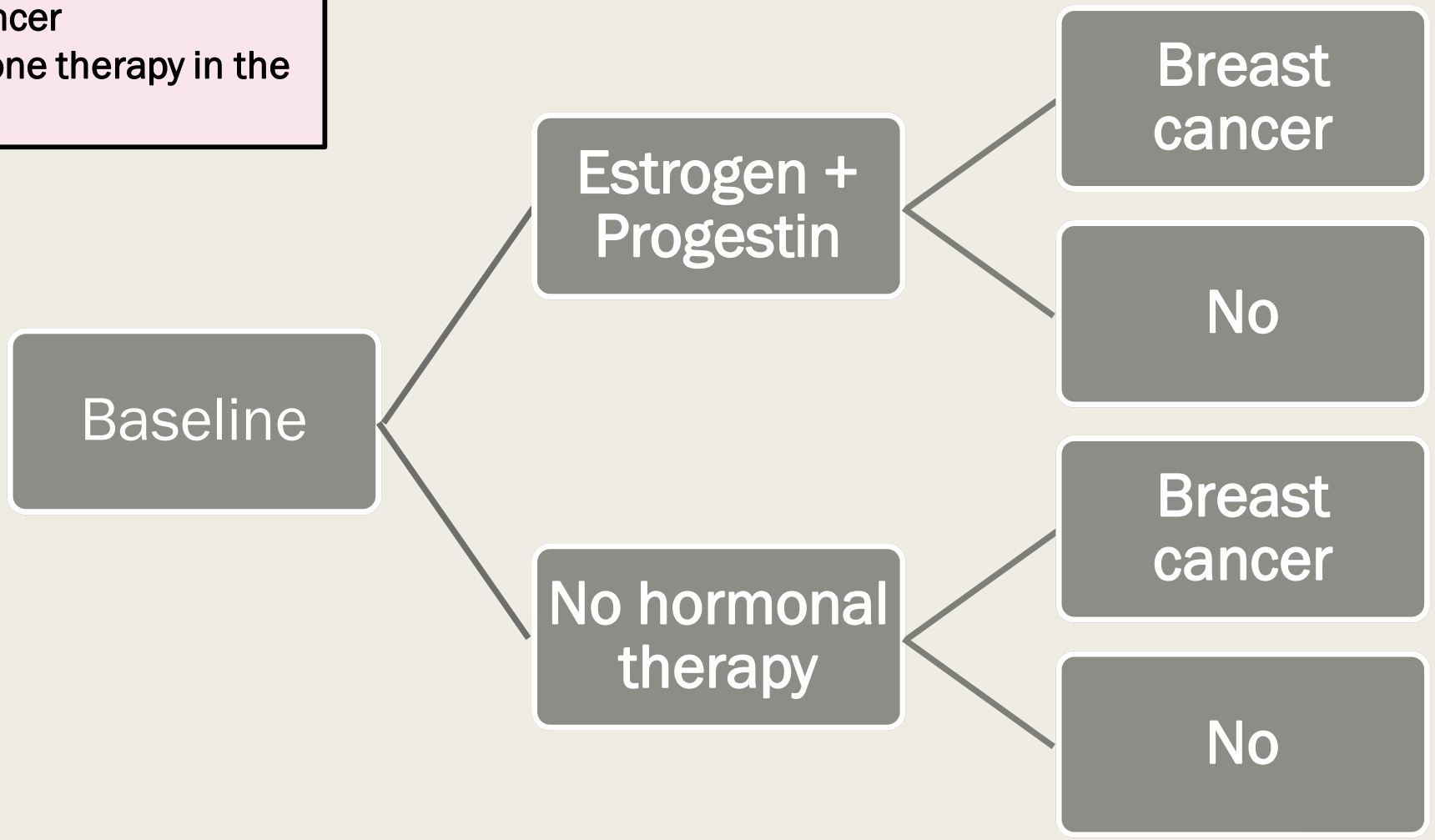
Table 1. A Summary of the Protocol of a Target Trial to Estimate the Effect of Postmenopausal Hormone Therapy on the 5-Year Risk of Breast Cancer

Protocol Component	Description
Eligibility criteria	Postmenopausal women within 5 years of menopause between the years 2005 and 2010 and with no history of cancer and no use of hormone therapy in the past 2 years.
Treatment strategies	Refrain from taking hormone therapy during the follow-up. Initiate estrogen plus progestin hormone therapy at baseline and remain on it during the follow-up unless you are diagnosed with deep vein thrombosis, pulmonary embolism, myocardial infarction, or cancer.
Assignment procedures	Participants will be randomly assigned to either strategy at baseline and will be aware of the strategy to which they have been assigned.
Follow-up period	Starts at randomization and ends at diagnosis of breast cancer, death, loss to follow-up, or 5 years after baseline, whichever occurs first.
Outcome	Breast cancer diagnosed by an oncologist within 5 years of baseline.
Causal contrasts of interest	Intention-to-treat effect, per-protocol effect
Analysis plan	Intention-to-treat effect estimated via comparison of 5-year cancer risks among individuals assigned to each treatment strategy. Per-protocol effect estimation requires adjustments for pre- and postbaseline prognostic factors associated with adherence to the strategies of interest. All analyses will be adjusted for pre- and postbaseline prognostic factors associated with loss to follow-up (57). This analysis plan implies that the investigators prespecify and collect data on the adjustment factors.

Eligibility criteria

1. Postmenopausal women within 5 years of menopause between the years 2005 and 2010
2. No history of cancer
3. No use of hormone therapy in the past 2 years

Study flow



Prior to proceeding with the emulation of the target trial: Data validation studies are required

- Because the data were not collected for research purposes, data codes may be inconsistent or ambiguous. Hence, researchers unfamiliar with the data must both consult with knowledgeable data users and **conduct data validation studies**

Eligibility criteria

- **Key: Our observational analysis should apply the same eligibility criteria used in the target trial**
- Suppose we want to emulate a randomized trial in which individuals will be followed via their contacts with the health care system. This target trial would only include individuals who can be expected to remain actively engaged with their health care providers during the follow-up period.

  A common strategy to emulate this criterion is to

  - *Restrict the analysis to individuals who have been in regular contact with the health care system before baseline (e.g., thos… ups or filled any prescriptions within the 2 pre… will remain in contact thereafter*
  - ***Note that we cannot simply exclude individuals whose claims are no longer found in the database some time after baseline. Rather, we must regard such individuals as lost to follow up (i.e., censored).***

** Keep in mind at all times **
We are doing the RCT (pretended)
We do not know the future

Treatment strategies

- **Fact**: We cannot emulate a placebo-controlled trial with **tight monitoring and enforcement of adherence to the study protocol**
- **Strategy coding** for eligible participants
  - ***First strategy****: eligible women who did not start hormone therapy*
  - ***Second strategy****: eligible women who did start estrogen plus progestin therapy*
  - *Otherwise eligible individuals who did not start any of the strategies of interest are* ***considered ineligible for the target trial emulation*** *and* ***excluded*** *from the observational analysis (****women who started estrogen only therapy will not participate in the emulation even if they meet all of the eligibility criteria****)*

Assignment procedures

- **Blinding**:
  - *We can only emulate target trials **without blind assignment**, which is the standard design of pragmatic trials, because **individuals in the data set and their health care workers are usually aware of the treatments that participants receive***

- **Randomization**:
  - *To emulate the random assignment of strategies at baseline, we need to **adjust for all confounding factors required to ensure comparability** (exchangeability) of the groups defined by initiation of the treatment strategies*

Assignment procedures:

Adjustment for confounders

- May be performed via
  - *Matching (perhaps on the propensity score)*
  - *Stratification or regression*
  - *Standardization or inverse probability weighting, g-estimation,*
  - *Doubly robust methods*

"If the observational database does not contain sufficient information on baseline confounders or if we fail to identify them, successful emulation of the target trial's random assignment is not possible"

Well-designed database is required to include all potential confounders

Assignment procedures:

Look for unmeasured confounding since it may be the cause of emulation failure

- Although it is generally impossible to determine whether the emulation failed because of uncontrolled confounding, indirect approaches may alert about possible **unmeasured confounding**
  - ***"Reversed" strategies**: a trial in which hormone therapy users are assigned to the strategies of "continue using therapy" or "stop using therapy".*
    - **Incompatible or surprising effect estimates** (e.g., a decreased risk both when initiating therapy in our original target trial and when discontinuing therapy in the reversed target trial) suggest that at least 1 of the 2 emulations failed to ensure a fair comparison.
  - ***Consider outcome controls for which no causal effect is expected**. If the confounders for the study and control outcomes are sufficiently similar, then the use of outcome controls can help detect confounding*
  - ***Machine-learning tools and other computer science techniques** might also help investigators search for combinations of variables that improve confounding adjustment compared with traditional methods*

Outcome

- Independent outcome validation is often warranted, because several studies have shown that lack of outcome validation may result in misleading effect estimates
- We often would prefer to emulate a target trial with systematic and blind ascertainment of the outcome to ensure that knowledge of treatment status does not influence a doctor's decision to look for the outcome
- Nonetheless, because doctors will generally be aware of the treatment received by the individual, we cannot use observational data to emulate a target trial with systematic and blind outcome ascertainment except when outcome ascertainment cannot be affected by treatment history (e.g., if the outcome is death and is independently ascertained from a death registry)

Causal contrast(s) of interest

- If the intention-to-treat and per-protocol effects are of interest in the target trial, we would try to estimate analogs of both effects from our observational data
  - *Intention-to-treat effect* *(i.e., the comparative effect of being assigned to the treatment strategies at baseline, regardless of whether the individuals continue following the strategies after baseline)*
  - *Per-protocol effect* *(i.e., the comparative effect of following the treatment strategies specified in the study protocol)*

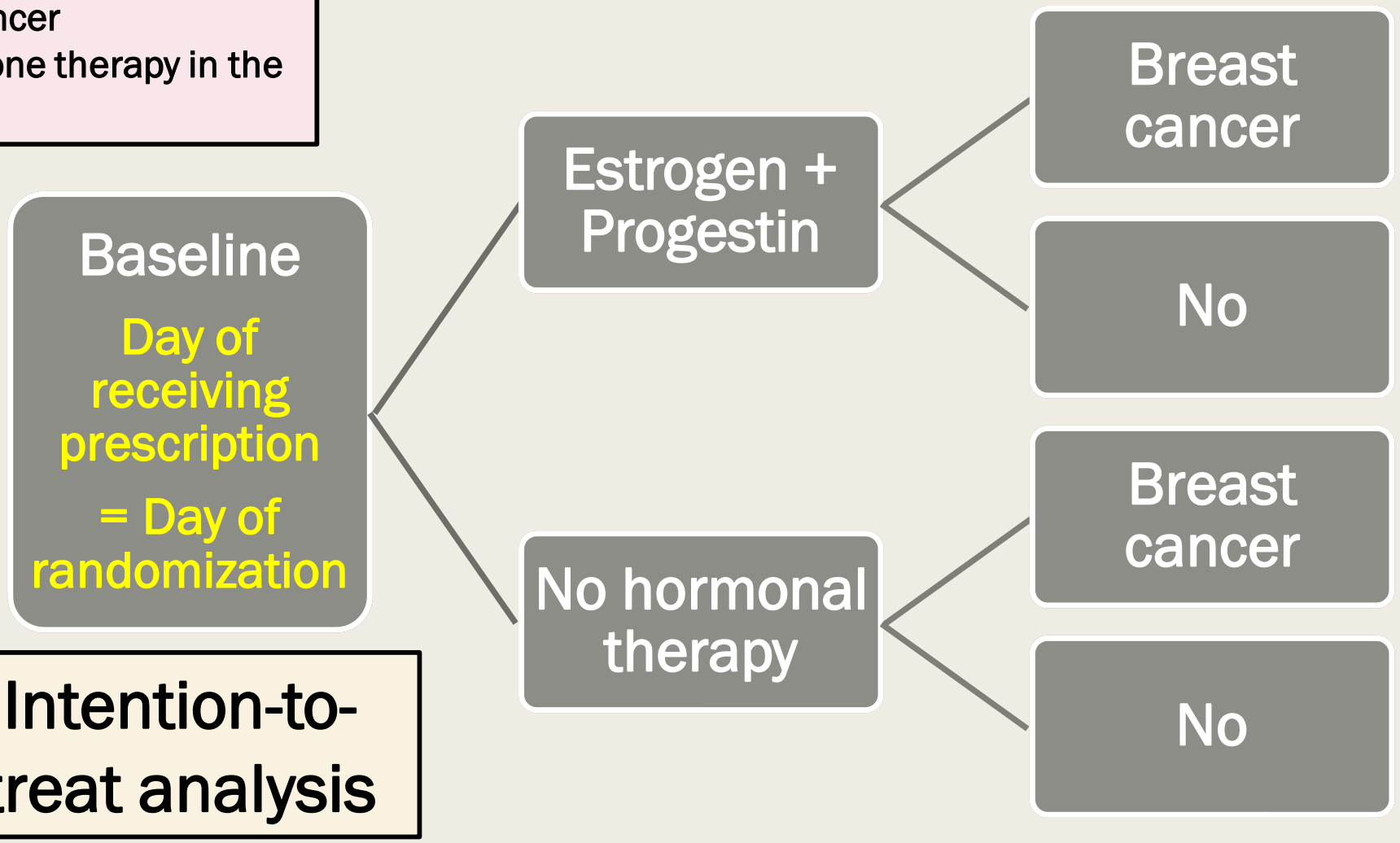
Analysis plan

- An intention-to-treat analysis, however, is rarely possible in observational analyses of existing data.
- The closest observational analog of the intention-to-treat analysis is a comparison of initiators of the different treatment strategies, assuming adequate adjustment for baseline confounders
  - *if we had data on prescription (rather than dispensing) of therapy, a **comparison of groups according to whether they did or did not receive a prescription of therapy at baseline** would be somewhat more analogous to the intention-to-treat analysis in the target trial*

Eligibility criteria

1. Postmenopausal women within 5 years of menopause between the years 2005 and 2010
2. No history of cancer
3. No use of hormone therapy in the past 2 years

Study flow



Intention-to-treat analysis

Analysis plan

- To estimate the per-protocol effect in both true randomized trials and emulated trials like ours, adjustment for baseline and post-baseline confounding is necessary when the treatment strategies under study are sustained over time. Because post-baseline prognostic factors associated with subsequent adherence to the strategies may be affected by prior adherence

- In the presence of selection bias due to loss to follow-up, adjustment for post-baseline factors might also be needed to validly estimate both intention-to-treat effects and per-protocol effects in both actual trials and observational analyses that emulate a target trial. When the post-baseline adjustment factors are affected by the treatment strategies themselves, g-methods are generally needed

G-estimation

- Observational studies of **the effect of a time-varying exposure** are nearly always plagued (= suffered) by **time-varying confounding**
- This phenomenon is caused by time-varying prognostic factors of the outcome that influence the exposure at each measurement time, and thereby distort the association between exposure and outcome over time
  - ❖ *For instance, the association between **physical activity** and **functional performance** (e. g. walking speed) in subjects with radiographic knee osteoarthritis may be **confounded by the extent of knee pain and symptoms** (Mansournia et al. 2012).*
    - ✓ Adjustment for such prognostic factors is usually difficult because they may themselves be influenced by previous exposures. For instance, the extent of knee pain may be influenced by the history of physical activity. When that happens, standard regression methods to adjust for confounding are fallible because they employ one and the same model to infer the effect of early versus late exposures; they thereby cannot avoid undue control for time-varying prognostic factors that are intermediate on the causal pathway from early exposure to later outcome (Robins 1986, 2000).

Vansteelandt S and Sjolander A; Epidemiol. Methods 2016; 5(1): 37–56

G-estimation

- In observational studies, a common goal is to estimate the effect of a point **exposure A** on an end-of-study **outcome Y** while accounting for a set of **confounders L**.
- When the set of confounders is large, stratification on each level l of L is no longer possible, and one must resort to modeling
  - *One option is **to fit an outcome regression model**, with Y as the dependent variable and A and L as covariates. Subsequent inference on the exposure effect then relies on having correctly modeled not only the association between Y and A but also the association between Y and L.*
  - *An alternative approach is **to model the relationship between the exposure and confounders (e.g. G-estimation)***
- **Robins' generalized methods (g methods)** provide consistent estimates of contrasts (e.g. differences, ratios) of potential outcomes under a less restrictive set of identification conditions than do standard regression methods (e.g. linear, logistic, Cox regression).

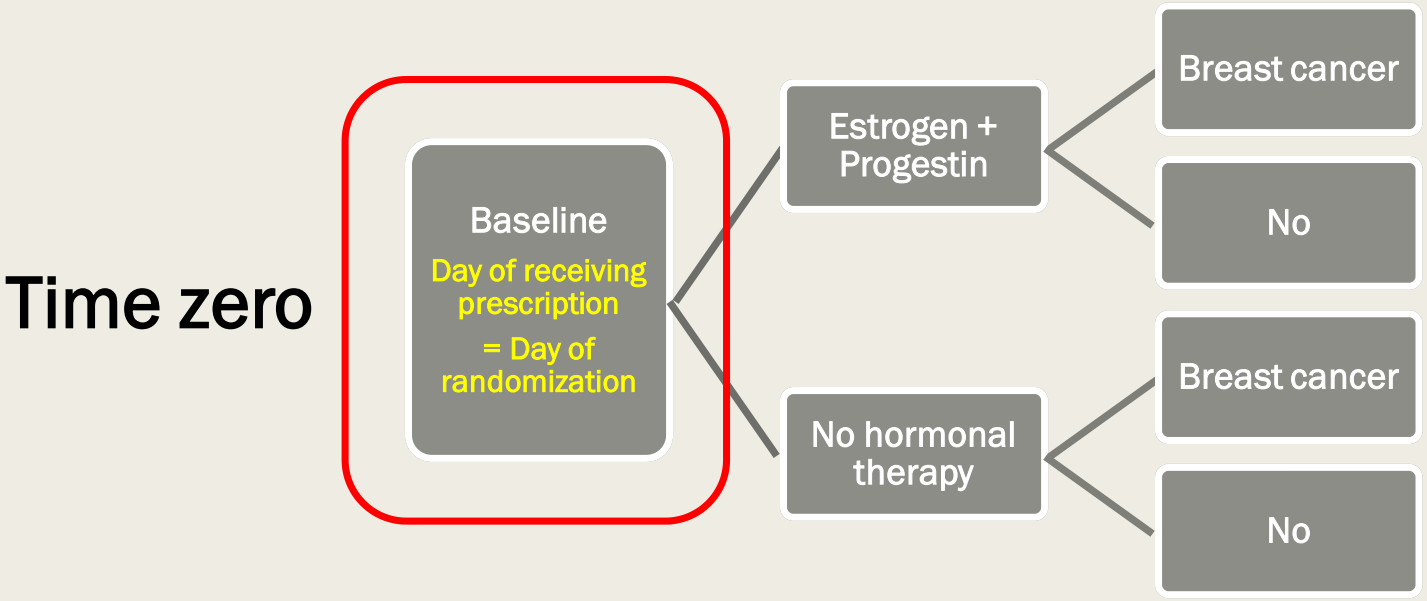
Dukes O and Vansteelandt S; Am J Epidemiol 2016; 187(5); 1079-1084
Naimi AI et.al.; Int J Epidemiol 2017; 756-762

Defining time zero

- "Time zero of follow-up" or "Baseline"

  "Eligibility criteria need to be met at that point but not later; study outcomes begin to be counted after that point but not earlier"

- With observational data, the best way to emulate time zero of the target trial is to define time zero to be the time when an eligible individual initiates a treatment strategy



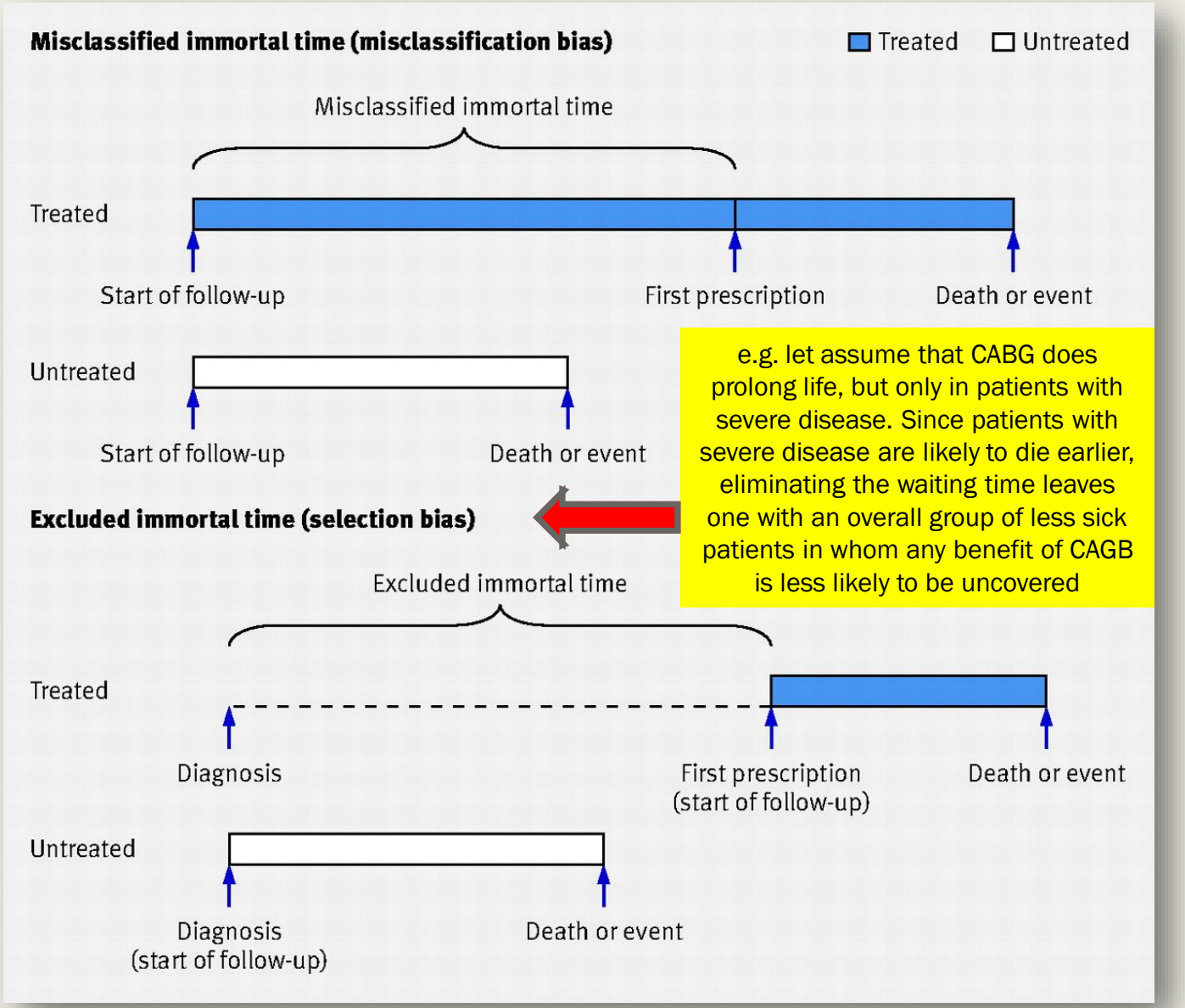
Specifying a grace period

- "Once a patient and her clinician decide that the patient should initiate hormone therapy, it may take several weeks to complete the clinical tests (e.g., a bone density scan and a lipid panel) and administrative procedures required before treatment initiation. Therefore, the trial protocol might specify that a women assigned to the strategy "initiate hormone therapy" is allowed a 1-month grace period so that she is considered compliant with the protocol if she initiates therapy within a month"
- In emulating a target trial that includes grace periods using observational data, we must allow for an analogous grace period measured from time zero. The use of a target trial with a grace period not only ensures that the strategies remain realistic but also increases the number of people in the observational database whose data can be used to emulate the target trial
- Analyses with a grace period at baseline are geared towards estimating a per-protocol effect of a target trial

Immortal bias

alias survivor (ship) bias

- ✓ Immortal time refers to a period of follow-up during which, by design, death or the study outcome cannot occur
- ✓ **Patients are immortal by definition before exposure**
- ✓ Immortal time typically arises when the determination of an individual's treatment status involves a delay or wait period during which follow-up time is accrued—for example, waiting for a prescription to be dispensed after discharge from hospital when the discharge date represents the start of follow-up



Lévesque LE et al.; BMJ 2010; Ho AMH; Anesthesia 2013

Discussion

- The target trial approach is consistent with a formal counterfactual theory of causality (e.g., new-users design, negative outcome controls)

1
- Specification of the protocol of the target trial
- Typically be an iterative process during which we will learn which particular target trials may be reasonably supported by the observational data

2
- Choose the one that is closest to the ideal trial that we would have liked to conduct to answer our question

3
- Outline a protocol, present a flow chart, summarize how the observational data set is used to emulate the target trial

4
- Explain how the target trial differs from the ideal trial

Discussion

- An explicit target trial approach is also advantageous to improve the quality of big data
- When investigators can influence how data are being actually recorded, a **target trial approach helps them identify critical data items for comparative effectiveness research** and **articulate a compelling rationale to modify data structuring or recording practices**
- When investigators from different institutions use a Common Data Model, an **explicit target trial approach may assist them in the development and evolution of the structure and contents of their data model**

Discussion

- The target trial approach allows us to systematically articulate the tradeoffs that we are willing to accept. **This explicit approach, in combination with subject-matter expertise, epidemiologic and methodologic proficiency, and innovative computer science tools**, seems our best bet to maximize the societal benefits of big data for causal inference

THANK YOU FOR YOUR ATTENTION