With Big Data Comes Big Responsibility

Using health care data to emulate randomized trials when randomized trials are not available

Miguel A. Hernán
Departments of Epidemiology and Biostatistics
Harvard School of Public Health

Seattle Symposium on Healthcare Data Analytics, September 2014

Health care: We need to make decisions NOW
- Treat with A or with B? Treat now or later? When to switch to C?
- A relevant randomized trial would, in principle, answer each comparative effectiveness and safety question
  - Interference/scaling up issues aside

But we rarely have randomized trials
- expensive, untimely, unethical, impractical
- And deferring decisions is not an option
  - no decision is a decision: “Keep status quo for now”
- Question: What do we do?

Answer: We conduct observational studies
- but only because we cannot conduct a randomized trial
- Observational studies are not our preferred choice
  - For each observational study, we can imagine a hypothetical randomized trial that we would prefer to conduct
  - If only it were possible

Effect of estrogen plus progestin hormone therapy on the 5-year risk of breast cancer among postmenopausal women
- Observational follow-up study
  - Women within five years of menopause, with no history of cancer, and who have not used hormone therapy for at least two years
  - Compare those who initiate and do not initiate hormone therapy at baseline
  - Identify those who receive a diagnosis of breast cancer over the next 5 years

Effect of estrogen plus progestin hormone therapy on the 5-year risk of breast cancer among postmenopausal women
- Open label, parallel randomized trial
  - Same
    - Except that therapy is randomly assigned at baseline
      - no other necessary differences between randomized and observational studies
  - Aside: In both observational and randomized studies, some women
    - discontinue or start hormone therapy after baseline
    - are lost to follow-up before the study ends.
The target trial

- An observational follow-up study can be viewed as an attempt to emulate a hypothetical, nonblinded randomized trial.
- If the observational study succeeds at emulating the target trial, both studies would yield identical effect estimates except for random variability.

Our proposed strategy for each clinical/policy question

- Step #1
  - Describe the protocol of the target trial.
- Step #2
  - Option A. Conduct the target trial.
  - Option B. Use observational data to emulate the trial.
  - Not perfect, but any better ideas anyone?

Enter Big Data

- A fashionable term for “observational data on many people”
  - e.g., large health care databases
- Epidemiologists, statisticians
  - have worked with big data for a long time and have learned to be cautious
  - now watch in disbelief how others promise the moon
  - Can we really use big data to understand everything?

Big Data for comparative effectiveness/safety research

- Better than Small Data
- But, regardless of size, observational data needs to emulate a target trial, which requires
  - Sufficient (and high-quality) longitudinal data are available for each individual
  - Appropriate emulation procedures are followed

Key elements of the protocol of the target trial

- Eligibility criteria
- Start/End of follow-up
- Strategies/Interventions
  - randomly assigned at start of follow-up
- Outcomes
- Causal effects of interest
  - e.g., intention-to-treat, per-protocol
- Analysis plan

The observational study needs to emulate

- Eligibility criteria
- Start/End of follow-up
- Strategies/Interventions
  - randomly assigned at start of follow-up
- Outcomes
- Causal effects of interest
  - e.g., intention-to-treat, per-protocol
- Analysis plan
Toy example

- Target trial:
  - HIV-infected patients who meet certain eligibility criteria at baseline and are followed for 5 years
  - Treatment groups: initiation of antiretroviral therapy A or antiretroviral therapy B at baseline
  - Outcome: all-cause mortality
  - Analysis: comparison of 5-year mortality risk for initiation of A versus B
- Observational study:
  - Same, except that interventions A and B are not randomly assigned

The target trial is a compromise

- between the trial we would really like to conduct and the trial we have any chances of emulating using the available data
- The drafting of the protocol of the target trial is typically an iterative process
- Good idea to add negative controls to the protocol
  - control exposures and outcomes

Yet emulation may not be straightforward

- Eligibility criteria
- Start/End of follow-up
- Strategies/Interventions
  - randomly assigned at start of follow-up
- Outcome
- Causal effect of interest
  - e.g., intention-to-treat, per-protocol
- Analysis plan

Eligibility criteria

- There may be insufficient data to characterize individuals eligible for the target trial
- Example:
  - In true trials, baseline screening to exclude prevalent, perhaps subclinical, cases
  - No baseline screening in most observational datasets

Start of follow-up

- When is time zero?

  - In true trials
    - the time of randomization
  - In emulated trials
    - ???

  - Failure to assign time zero correctly may lead to misunderstandings
    - e.g., immortal time bias, hormone therapy confusion

Definition of strategies/interventions

- In true trials
  - Relatively well-defined interventions in protocol, including start and end
- In emulated trials
  - Need to be relatively well-defined too
  - For example, defining treatment as “current” vs. “never” use does not generally correspond to well-defined interventions in a target trial
Randomized assignment of interventions

- In true trials
  - Expected
- In emulated trials
  - Emulation correct only if no unmeasured confounders for treatment assignment at baseline
  - Which cannot be guaranteed

Outcome

- In true trials
  - Targeted ascertainment
  - Blinded ascertainment
- In emulated trials
  - Ascertainment via routine care
  - Ascertainment may be affected by interventions themselves

Examples of trial emulation using Big Data

1. Electronic medical records - THIN
   - Statins and coronary heart disease
   - Static strategies
   - Treat vs. no treat
2. Claims database - USRDS Medicare
   - Epoetin and mortality
   - Static and Dynamic strategies
   - Intervention depends on response to previous intervention

EXAMPLE #1
Statins and heart disease

- Question
  - What is the effect of statin therapy on the risk of coronary heart disease?
- Extreme example of confounding
- Data: UK THIN (electronic medical records)
  - ~75,000 eligible patients
  - Used to emulate a sequence of observational “trials” of statin initiation
  - Danaei et al. Statistical Methods in Medical Research 2013

The target trial

- Eligibility criteria
  - Age 55–84, no history of cardiovascular disease and serious chronic diseases, no statin use within 2 years of baseline
- Strategies
  - Initiation of statin therapy
  - Standard of care without statin therapy
- Follow-up
  - From baseline until CHD, death, loss to follow-up or administrative end of the study
- Outcome
  - CHD

Observational “trials”

- Same eligibility criteria, treatment groups, and follow-up
- We emulated a sequence of 83 trials
  - Each month between Jan 2000 and Nov 2006 a new trial starts
  - A method to establish time zero
- Individuals may participate in more than one trial if they meet eligibility criteria
  - Generalization of new-users design
Flowchart of emulated “trials”

Adherence to treatment

Statistical analysis

- Observational analog of “intention to treat” effect
  - Cox model for statin initiation at baseline (yes, no) with baseline confounders as covariates
  - Use of propensity score yielded the same estimates, as expected
- Observational analog of “per protocol” effect
  - Cox model for statin initiation with baseline confounders as covariates
  - Artificial censoring after deviating from baseline treatment, i.e., initiating statins for non-initiator, stopping statins if initiator
  - Adjusted for time-varying confounders via IP weighting
- Potential confounders
  - Sex, age, LDL-cholesterol, HDL-cholesterol, BMI, smoking, alcohol use, systolic blood pressure, diabetes, hypertension, atrial fibrillation, use of antihypertensives, insulin, other lipid-lowering drugs, and beta-blockers, doctors visits, referrals, hospitalizations in last 3 months, etc.

Hazard ratio (95% CI) of CHD THIN “trials” 2000-2006

<table>
<thead>
<tr>
<th></th>
<th>Intention-to-treat effect</th>
<th>Per-protocol effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unique cases</td>
<td>633</td>
<td>438</td>
</tr>
<tr>
<td>Unique persons</td>
<td>74,806</td>
<td>74,806</td>
</tr>
<tr>
<td>Cases</td>
<td>6,335</td>
<td>4,849</td>
</tr>
<tr>
<td>Person-“trials”</td>
<td>844,800</td>
<td>844,800</td>
</tr>
<tr>
<td>Age-sex adjusted</td>
<td>1.29 (1.06, 1.56)</td>
<td>1.54 (1.09, 2.18)</td>
</tr>
<tr>
<td>Adjusted for covariates</td>
<td>0.89 (0.73, 1.09)</td>
<td>0.84 (0.54, 1.30)</td>
</tr>
<tr>
<td>Adjusted for covariates (excluding first year of follow-up)</td>
<td>0.71 (0.53, 0.94)</td>
<td>0.53 (0.27, 1.02)</td>
</tr>
</tbody>
</table>

What if we had compared prevalent (not incident) users vs. nonusers?

- Current users
  - HR: 1.42 (1.16, 1.73)
- Persistent (1 yr) current users
  - HR: 1.05
- Persistent (2 yrs) current users
  - HR: 0.77 (0.51, 1.18)
  - Confounding-selection bias tradeoff

Mortality hazard ratio for statins in CHD secondary prevention studies

- RCTs: 0.84 (0.77, 0.91)
- Observational studies
  - Incident users: 0.77 (0.65, 0.91)
  - Prevalent-incident mix: 0.70 (0.64, 0.78)
  - Prevalent users: 0.54 (0.45, 0.66)

10/4/2014 Hernan - Emulating Trials
**Other static comparisons: same analytic approach**

- **Head-to-head comparisons:**
  - Example: Lipophilic statins (atorvastatin, simvastatin) vs other statins
  - Danaei et al. *Diabetes Care* 2013
- **Joint interventions**
  - Example: statins plus antihypertensives vs other combinations
  - Danaei et al. 2014 (under review)

**EXAMPLE #2**

**Epoetin dosing and mortality**

- **Question:** What is the effect of different doses of epoetin therapy on the mortality risk of patients undergoing hemodialysis?
- **Data:** US Renal Data System (Medicare claims database)
  - ~18,000 eligible elderly patients

**The target trial**

- **Eligibility criteria**
  - End-stage renal disease
- **Strategies**
  - Fixed weekly dose of intravenous epoetin
  - 15,000, 30,000, or 45,000 units
- **Follow-up**
  - From 3 months after hemodialysis onset until death, loss to follow-up or administrative end of the study (1 year)
- **Outcome**
  - All-cause mortality

**Methodological challenge**

- **Time-varying treatment**
  - Use and dose of epoetin varies over the course of the disease
- **Time-varying confounders**
  - Hematocrit level, comorbidities
  - may be affected by prior treatment
- **Treatment-confounder feedback**
  - Need “causal” methods
  - Inverse probability weighting of marginal structural models

**Survival under 3 epoetin dosing regimes**


**But this is a silly target trial**

- In clinical practice, patients do not receive a fixed weekly dose of epoetin
  - That would be clinical malpractice
- Rather, actual clinical strategies are dynamic
  - A patient’s weekly dose depends on her hemoglobin or hematocrit, which in turn depends on her prior weekly dose
More reasonable strategies for a target trial

1. Mid-Hematocrit strategy
   - epoetin to maintain Hct between 34.5% and 39.0%

2. Low-Hematocrit strategy
   - epoetin to maintain Hct between 30.0% and 34.5%

- Under both strategies, epoetin dose is
  - increased by >10% if previous Hct below target
  - decreased by <10% times [previous Hct minus lower end of range] or increased by <10% times [upper end of range minus Hct] if Hct within target
  - decreased by >25% if Hct above target

More reasonable strategies imply more work

- Need to specify a more detailed protocol for the target trial
- Need to specify how to emulate that protocol
  - Appropriate adjustment for time-varying confounders becomes critical
  - Zhang et al. Medical Care 2014

Survival under these 2 dynamic strategies
A common misinterpretation

☐ You are saying that observational studies are as good as RCTs?
    ■ "This is a cohort study that tries to turn itself into a clinical trial. This involves a series of assumptions and manoeuvres which lack credibility."
    ☑ Anonymous JAMA reviewer, April 2014

☐ No, the point is not that observational studies can turn themselves into randomized experiments
    ■ They can’t

The point is that we can do better

☐ by using observational data to emulate randomized trials

☐ The limitations of observational studies (e.g., confounding, mismeasurement) remain, but we do not compound them with additional biases

Remember

☐ Observational studies are what we do when we cannot conduct a randomized trial
    ■ In the absence of practical and ethical constraints, sane people will always prefer a randomized trial

The target trial in comparative effectiveness/safety research

☐ Unifying concept
    ■ Can be applied to all types of designs for causal inference about the effects of interventions
    ■ Randomized and non-randomized

☐ Organizing principle
    ■ Puts together causal inference concepts/methods dispersed throughout the literature

The benefits of being explicit:
Using the “target trial” helps

☐ provide constructive criticism of observational studies
    ■ which components of the target trial we weren’t able to mimic approximately?
    ■ which components of the target trial would be problematic even if we were able to conduct a true trial?

☐ understand why estimates differ across studies
    ■ assess the sensitivity of estimates to different design choices for the target trial
    ■ focus research efforts on “sensitive” choices

☐ avoid some methodologic pitfalls

If we want to know whether observational studies “work”

☐ We first need to know what question is being asked exactly
    ■ Being explicit requires a detailed description of the target trial

☐ Only then can we discuss the design and analysis of the observational data
    ■ That is, we call the biostatisticians after the clinicians
No alternative to observational studies

- So we better keep improving them
  - Because people will keep using (Big and Small) data to guide their decisions

- Present challenge: combining the concept of target trial with data mining technologies
  - i.e., combining subject-matter knowledge with automatic procedures