High Dimensional Propensity Scores in an Automated Setting

Jeremy A. Rassen
May 2013

Div. of Pharmacoepidemiology
Dept. of Medicine
Brigham & Women’s Hospital
Harvard Medical School
Four non-randomized studies

Log(Risk Ratio)

Unadjusted
Age/Sex Adjusted
“Fully Adjusted”

... what now?
Some approaches

- Either remove confounders (restriction, etc.) or remove the effect of confounding.
- Use approaches to deal with unmeasurable confounding
  - Instrumental variables
  - External adjustment
  - [Trials]
- Measure some previously unmeasured confounders
  - Where to get the data?
Search all the data available for confounders proxies

- Measured confounders (such as age) serve as proxies for unmeasured factors (such as general state of health).

Diagram:
- Comorbidity
- Age
- Treatment
- Outcome
Other examples of proxies

<table>
<thead>
<tr>
<th>Observable Proxy</th>
<th>Implied Health State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of oxygen canister</td>
<td>Very frail</td>
</tr>
<tr>
<td>Hypertension diagnosis during hospital stay</td>
<td>Not too bad off</td>
</tr>
<tr>
<td>Annual checkup and colonoscopy</td>
<td>Careful with health; compliant patient</td>
</tr>
<tr>
<td>Incident statin use at age 70</td>
<td>Fairly healthy</td>
</tr>
<tr>
<td>Many drugs used, many office visits</td>
<td>Fairly ill</td>
</tr>
</tbody>
</table>
An uncomfortable thought?

- It may not matter what you’re measuring as long as you can tell whether you’re measuring something “important”.
Automated variable identification

- Identify variables that appear to be confounders, even if they’re only proxies.
- Measure as many of these as possible and adjust for them.
- A technical problem: you can’t adjust for too many variables in a model (roughly 1:10 per outcome event)
Propensity scores and proxies

The approach:
- Collect as many codes as possible
- Identify those codes that could possibly bias the exposure/outcome relationship
- Combine variables identified *a priori* with the “best” of these codes in a propensity score.
- Use this “high dimensional propensity score” to adjust for confounding.

Currently implemented in a SAS macro or R program.
Where can new variables come from?

**Data type**
- Inpatient Diagnoses *
- Outpatient Diagnoses *
- Inpatient Procedures **
- Outpatient Procedures **
- Medication dispensings ***
- Lab test results
- Unstructured text notes

**Frequency/Intensity**
- Once
- Sporadic
- Frequent

**Temporality**
- Proximal to exposure
- Evenly distributed
- Distal to exposure start
High-dimensional propensity score

Prevalence of factors

Frequency, temporal clustering

Interactions

Basic covariate prioritization re confounding

Boost through DRS machine learning

PS estimation followed by matching, stratification

Target parameter estimation for causal inference
hd-PS algorithm (1/7)

- **STEP 1a. Specify Data Sources**
  Define $p$ data dimensions; use data stream of 180 days up to the initiation of study exposure. Collect all codes from each of the $p$ dimensions.

- **STEP 1b.**
  Include basic demographic information (age, sex) and investigator-defined covariates (history of diabetes, use of statins, ...)

CONFIDENTIAL
STEP 2. Identify empirical candidate covariates
Within each data dimension sort by prevalence of codes. Identify the $n$ most prevalent codes and how the codes should be treated.

Example:
- $n=200$
- Use 3 digit ICD-9 (vs. 4 or 5 digit)
- Use 5 digit CPT
- Use generic drug name (vs. drug class)
STEP 3. Assess code recurrence and create indicator variables for each patient

For each identified code, address frequency by creating 3 variables:

- CovX_once = 1 if that code appeared at least once within 180 days
- CovX_sporadic = 1 if code appeared at least more than the median
- CovX_frequent = 1 if code appeared at least more than the 75th percentile.
STEP 4. Prioritize covariates
Calculate for each covariate the possible amount of confounding the covariate could adjust for.

Use the Bross (1966) formula:

$$BIAS_{mult} = \frac{P_{C1}(RR_{CY} - 1) + 1}{P_{C0}(RR_{CY} - 1) + 1}$$
STEP 5. Select covariates
Add the demographic covariates and investigator-defined covariates from step 1. Then, select top $k$ empirical covariates from step 4 as ranked by potential bias.

Example:
- 4 demographic covariates
- 24 investigator-defined covariates
- $k=500$ hd-PS-selected covariates
hd-PS algorithm (6 + 7/7)

- **STEP 6. Estimate propensity score (the hd-PS)**
  Estimate propensity score using multivariate logistic regression, including all investigator-defined covariates and the k hd-PS-selected covariates.

- **STEP 7. Use the hd-PS**
  Use the hd-PS in an outcome model as any PS would be used (matching, deciling, trimming, etc.)
Recall the four studies
Recall the four studies
Recall the four studies
Data sources

Insurance claims data:
U.S. Medicare
U.S. commercial
Canada
Germany
Elect. health records:
United Kingdom
Regenstrif

(b) Schneeweiss S, Rassen JR, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology 2009;20:512–22.
(c) Patorno E, Bohn RL, Wahl PM, Avorn J, Patrick AR, Liu J, Schneeweiss S. Anticonvulsant medications and the risk of suicide, attempted suicide, or violent death. JAMA 2010;303:1401-9
(d) Same as (b)
hdPS simulation performance

Simulated true effect estimate* (HR= 1.0, 1.5, 2.0)

* Plasmode simulations inject a defined causal effect of $E$ on $Y|C$ in a given healthcare database preserving the underlying data structure and information content.
Getting the software

- All materials available for download at: http://www.drugepi.org

- Open source (Mozilla 2.0 license)
### Some more examples

<table>
<thead>
<tr>
<th>Exposure (referent)</th>
<th>Outcome</th>
<th>Relative risk estimate (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted</td>
<td>Adjusted by basic variables</td>
</tr>
<tr>
<td><strong>Coxibs (ns-NSAIDS)</strong>&lt;sup&gt;7&lt;/sup&gt;</td>
<td>GI bleed within 180 days</td>
<td>1.09 (0.91–1.30)</td>
</tr>
<tr>
<td><strong>Coxibs (ns-NSAIDS)</strong>&lt;sup&gt;41&lt;/sup&gt;</td>
<td>GI bleed</td>
<td>1.21 (0.91–1.61)</td>
</tr>
<tr>
<td><strong>Coxibs (ns-NSAIDS)</strong>&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Upper GI bleed</td>
<td>1.50 (0.98–2.28)</td>
</tr>
<tr>
<td><strong>Statins (glaucoma drugs)</strong>&lt;sup&gt;7&lt;/sup&gt;</td>
<td>All-cause mortality within 180 days</td>
<td>0.56 (0.51–0.62)</td>
</tr>
<tr>
<td><strong>TCAs (SSRIs), &lt;18 y.o.&lt;sup&gt;9&lt;/sup&gt;</strong></td>
<td>Suicide within 1 year</td>
<td>0.59 (0.28–1.27)</td>
</tr>
<tr>
<td><strong>TCA (SSRIs), 18+ y.o.&lt;sup&gt;10&lt;/sup&gt;</strong></td>
<td>Suicide within 1 year</td>
<td>0.97 (0.77–1.21)</td>
</tr>
<tr>
<td><strong>Clopidogrel + PPI (clopidogrel alone)</strong>&lt;sup&gt;42&lt;/sup&gt;</td>
<td>MI or CV death</td>
<td>1.74 (1.44–2.10)</td>
</tr>
<tr>
<td><strong>Neurontin (Topamax)</strong>&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Suicide or attempted suicide</td>
<td>0.95 (0.76–1.19)</td>
</tr>
<tr>
<td><strong>Conventional APMs (atypical APMs)</strong>&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Death</td>
<td>1.37 (1.11–1.69)</td>
</tr>
<tr>
<td><strong>Benzodiazepines (atypical APMs)</strong>&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Death</td>
<td>1.37 (1.14–1.64)</td>
</tr>
</tbody>
</table>
Using hd-PS in a sequential monitoring environment

I. Basic design choice and inputs

II. Semi-automated balancing & estimation

III. Data aggregation

IV. Alerting algorithm selection and application

V. Reporting
Drug of interest: rofecoxib

Comparators: non-selective NSAIDs

Other exposure parameters:
- New users (180-day washout)
- “As-treated” with 30-day maximum gap

Outcome of interest: AMI (inpatient ICD-0 410.x1)

Pre-defined covariates:
- 18 ICD-9 based conditions
- 9 drug classes

hdPS options: defaults + health service utilization variables

Databases: PACE and PAAD, 5/20/1999 to 12/31/2004
Goals of diagnostics:
- To identify errors and areas of missing data
- To assess level of balance achieved
- Evaluate and compare patient characteristics, numbers of events, rates across data partners
- To decide whether to include data in ongoing monitoring

Table 1: demographics, predefined covariates, M-distance

Figure 1: PS distribution, c-statistics

Additional figures: summary and detailed balance figures

* All of these can be produced for unmatched, matched, or stratified cohorts
## Vioxx Table 1

### Table 1. Cohort of New Initiators of Rofecoxib and Non-Selective NSAID (Unmatched)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>rofecoxib (N = 9409, 100.0%)</th>
<th>nsaid (N = 9977, 100.0%)</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>9409 (100.0%)</td>
<td>9977 (100.0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Events While on Therapy</td>
<td>39 (0.4%)</td>
<td>15 (0.2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Person time at risk</td>
<td>59.9 (33.3)</td>
<td>46.4 (32.5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Patient Characteristics

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>rofecoxib</th>
<th>nsaid</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>60-70</td>
<td>1305 (13.9%)</td>
<td>1679 (16.8%)</td>
<td>-2.9</td>
<td>-0.082</td>
</tr>
<tr>
<td>70-80</td>
<td>3631 (38.6%)</td>
<td>3883 (38.9%)</td>
<td>-0.3</td>
<td>-0.007</td>
</tr>
<tr>
<td>80-90</td>
<td>3179 (33.8%)</td>
<td>2619 (26.3%)</td>
<td>7.5</td>
<td>0.164</td>
</tr>
<tr>
<td>90-100</td>
<td>580 (6.2%)</td>
<td>395 (4.0%)</td>
<td>2.2</td>
<td>0.101</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender (F)</th>
<th>rofecoxib (N = 7764, 82.5%)</th>
<th>nsaid (N = 7374, 73.9%)</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>76.3 (10.7)</td>
<td>73.1 (12.2)</td>
<td>3.2</td>
<td>3.2</td>
</tr>
</tbody>
</table>

### Recorded use of:

<table>
<thead>
<tr>
<th>Recorded use</th>
<th>rofecoxib (N)</th>
<th>nsaid (N)</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ace Inhibitors</td>
<td>1224 (13.0%)</td>
<td>1351 (13.5%)</td>
<td>-0.5</td>
<td>-0.016</td>
</tr>
<tr>
<td>ARB</td>
<td>567 (6.0%)</td>
<td>535 (5.4%)</td>
<td>0.6</td>
<td>0.029</td>
</tr>
<tr>
<td>Anticoagulants</td>
<td>548 (5.8%)</td>
<td>328 (3.3%)</td>
<td>2.5</td>
<td>0.122</td>
</tr>
</tbody>
</table>

### Covariate Balance

| Mahalanobis Distance | 0.131 |

**Primary Analysis**

**Additional predefined covariates**
Diagnostics: match quality

**UNMATCHED**

**hd-PS MATCHED**
Variable selection is often based on the strength of association with treatment. Will tend to select IVs (instrumental variables) or near-IVs, variables that strongly predict treatment but have a weak association with outcome. Balancing IVs and near-IVs may amplify any residual confounding bias\(^1,2\) in some cases.

\(^2\) Pearl, *Proceedings of the UAI 2010.*
Simulation studies showed that increases in bias due to adjusting for IVs were generally small (< 10%) except in extreme examples\(^1\).

When in doubt whether a covariate is an IV, include the covariate. Unmeasured confounding bias is worse than Z-bias.

\(^1\) Myers, Rassen, Gagne, et al., AJE, 2011;174(11):1213-1222.
In early experience with a new drug, a prescriber’s use is likely to evolve

- Need to re-estimate propensity scores after brief time periods to model changes in therapeutic choice

- Emergence of new evidence on efficacy and safety in specific patient groups can also induce changes over time in propensity scores.
Even for drugs that go on to be blockbusters, initial time periods after approval will have few users, limiting ability to select variables for and fit high-dimensional propensity scores.

Distinction of true risk factors from variables that influence treatment choice but not outcome (instruments) is challenging with infrequently used drugs.

Thank you.