Methods to control for confounding

- Introduction & Overview -

Nicolle M Gatto
18 February 2015
Learning Objectives

At the end of this confounding control overview, you will be able to:

• Understand how confounding arises in pharmacoepidemiology and the problem it poses for causal inference
• Differentiate (exposure probability) weighting approaches to confounding control, including propensity score control, matching, and stratification, and inverse probability weighting
• Relate these weighting approaches to randomization
• List key assumptions necessary for these approaches to estimate causal effects
• Distinguish between point treatment and time-dependent confounding
Outline

- Ideal comparisons for causal inference
- How confounding impedes causal inference
- The gold standard
- Mimicking ideal comparisons without randomization:
  - Typical multivariate regression
  - Propensity score methods
  - Inverse probability weighting (IPW)
- Practical considerations
- Take home messages
Aim of comparative pharmacoepidemiology studies

Does drug A protect against outcome Z?

\[ A \xrightarrow{?} Z \]

Does drug A cause outcome Y?

\[ A \xrightarrow{?} Y \]

Here, the exposure of interest, A, is a “point-treatment”
The essential problem

L1
  \rightarrow L2
    \rightarrow ...L_i
      \rightarrow Drug A

- Outside the context of study, many reasons people take one medication over another
The essential problem

- These reasons are often related to outcomes of interest
The essential problem

• Confounding: When risk factors for the disease are associated with use of the medication of interest
The essential problem

- For simplicity let’s consider L as the totality of confounding
- “Point-treatment” confounding
The ideal (hint: it’s not the RCT)

Target Population $N$

- $N \rightarrow Y_{+|A=1}$
- $N \rightarrow Y_{+|A=0}$

$\frac{Y_{+|A=1}}{N} - \frac{Y_{+|A=0}}{N} = \text{RD}_{\text{causal}}$
The ideal (hint: it’s not the RCT)

- We’ve defined $RD_{\text{causal}}$ (for now) as the effect – with all else equal - in the total population if everyone was exposed versus no one was exposed
- But at least one of these conditions is counter-to-fact
- Instead we use substitutes for the ideal risks
The real world substitute

In target Population N, people become A+ or A-

\[ \text{An observable estimate} = \frac{\text{RD}_{\text{crude}}}{(Y_{+ \text{in } A^+} / A^+)} - (Y_{+ \text{in } A^-} / A^-) \]
The real world substitute

In target Population N, people become A+ or A-

An observable estimate

$$\text{An observable estimate} = \text{RD}_{\text{crude}}$$

$$Y_{+\text{in A+}} = \frac{Y_{+\text{in A+}}}{[f_1 N]} - \frac{Y_{+\text{in A-}}}{A_-}$$
The real world substitute

In target Population N, people become A+ or A-

\[
\text{An observable estimate} \rightarrow \frac{Y_{+\text{in A+}}}{f_1N} - \frac{Y_{+\text{in A-}}}{(1-f_1)N} = \text{RD}_{\text{crude}}
\]
The real world substitute

In target Population N, people become A+ or A-

\[ \text{RD}_{\text{crude}} = \frac{\left(\frac{f_2 f_1 Y_{+in\ A+}}{f_1 N}\right)}{\left(\frac{Y_{+in\ A-}}{(1-f_1)N}\right)} \]

An observable estimate
The real world substitute

In target Population N, people become A+ or A-

\[ \frac{([f_2 f_1 Y_{+A+}] / [f_1 N])}{([f_3(1-f_1)Y_{+A-}] / [(1-f_1)N])} = RD_{\text{crude}} \]

An observable estimate
The real world substitute

In target Population N, people become A+ or A-.

The equation for the real world substitute is:

\[ RD_{\text{causal}} \neq RD_{\text{crude}} \]

With confounding by L.
Possible solutions

• Try to avoid confounding through design, e.g.
  - Randomize
  - Match
  - Restrict

• Try to control confounding through analysis, e.g.:
  - Traditional analysis methods
  - Techniques that weight by causes of the treatment (aka mimic randomization)
  - Techniques that weight by causes of the outcome
Why randomization is “the gold standard”

Target Population N, randomized 1:1 to A+ or A-

\[ \frac{Y_{+inA_+}}{A_+} - \frac{Y_{+inA_-}}{A_-} = RD_{crude} \]

An observable estimate
Why randomization is “the gold standard”

Target Population N, randomized 1:1 to A+ or A-

\[ \frac{(Y_{+ \text{in } A+})}{0.5N} - \frac{(Y_{+ \text{in } A-})}{A-} = RD_{\text{crude}} \]

An observable estimate
Why randomization is “the gold standard”

Target Population N, randomized 1:1 to A+ or A-

\[ \frac{0.5Y_{+in A+}}{0.5N} \]

\[ (Y_{+in A-} / A-) \]

= \[ \text{RD}_{\text{crude}} \]
Why randomization is “the gold standard”

Target Population N, randomized 1:1 to A+ or A-

\[ RD_{\text{causal}} = RD_{\text{crude}} \]

An observable estimate
Why randomization is “the gold standard”

Target Population N, randomized 1:3 to A+ or A-

\[ RD_{\text{causal}} = RD_{\text{crude}} \]
Why randomization is “the gold standard”

- If randomization is successful, no confounding at baseline
Possible solutions

- Try to avoid confounding through design, e.g.
  - Randomize
  - Match
  - Restrict

- Try to control confounding through analysis, e.g.:
  - Traditional analysis methods
  - Techniques that weight by causes of the treatment (aka mimic randomization)
  - Techniques that weight by causes of the outcome
Basic requirements of these methods

• Appropriately hypothesize factors that contribute to confounding
• Can measure these factors (validly)

...‘there is no unmeasured confounding’
Option 1: Traditional confounding control

- For example, multivariate regression analysis
- Estimates $RD_{\text{causal}}$? Yes, when:
  - All variables that contribute to confounding are included in the model
  - Statistical model is specified correctly
  - Effect of $A \rightarrow Y$ across strata of confounding variables is homogeneous
Option 2: Propensity score (PS) methods

- Estimates $RD_{\text{causal}}$? Yes, when:
  - All variables that contribute to confounding are included in the model
  - Statistical model is specified correctly
  - Effect of $A \rightarrow Y$ across strata of confounding variables is homogeneous
    - (to estimate $RD_{\text{causal}}$ as we’ve defined it: the total population effect)
Option 3: Inverse probability weighting (IPW)

- Estimates $\text{RD}_{\text{causal}}$? Yes, when:
  - All variables that contribute to confounding are included in the model
  - Statistical model is specified correctly
  - Positivity holds
    - There are no strata of $L$ for which there are no subjects who received a particular exposure level
To estimate $\text{RD}_{\text{causal}}$ in total population

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Traditional methods</th>
<th>Propensity score control</th>
<th>IPW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requires all confounding factors measured?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Statistical correctly specified</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Can handle large number of confounding variables?</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Requires homogeneity?</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Requires positivity?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Estimating the propensity score

- Traditional PS: probability of being exposed conditional on predictors of exposure
- Minimally sufficient PS: probability of being exposed conditional on variables that contribute to confounding
Estimating the propensity score

• Traditional PS: probability of being exposed conditional on predictors of exposure

• Minimally sufficient PS: probability of being exposed conditional on variables that contribute to confounding
  - A weight is assigned to each person according to the conditional probability of exposure given the confounder(s)
Estimating the propensity score: Example 1

Weight these people by \( \text{pr}(A^+ | L^+) \)
Estimating the propensity score: Example 2

Study Population (N)

L+

A+ → Y+
  → Y-
L-

A+ → Y+
  → Y-
A- → Y+
  → Y-

Weight these people by $\text{pr}(A+ \mid L-)$
Estimating the propensity score

Example SAS code:

```sas
**estimate propensity scores**;
proc logistic descending data=dataset1;
    model A=L1 L2 L4 L5 L7 L8;
    output out=probA1_PS p=prA1_PS  xbeta=logit_ps;
run;
```
Applying the propensity score

- Options: PS control, stratification & matching
- After estimating PS for each person, use it to control confounding by:
  - Adjusting for the PS as a covariate
  - Stratifying the effect estimate by the PS
  - Matching A+ and A- people by PS
Applying the propensity score

Example SAS code:

**PS adjusted (as covariate) RR**;
proc genmod data=probA1_PS;
   model Y=A prA1_PS / dist=poisson;
run;
Applying the propensity score

Example SAS code:

**PS quintile stratification for CMH RR**;
proc rank data=probA1_PS groups=5 out=tempPS;
   ranks rnks;
   var prA1_PS;
run;

proc sort data=tempPS;
   by A Y;
proc freq data=tempPS order=data;
   tables rnks*A*Y / cmh;
   output out=tempPS1 mhrrc2 lgrrc2;
run;

data PSstrat_5;
   set tempPS1;
   beta=log(1/_MHRRC2_);
   SE=(log(U_MHRRC2)-log(L_MHRRC2))/3.92;
run;
Estimating the inverse probability weight

- Another way to use the propensity score!
- IPW: Inverse of the exposure probability conditional on variables that contribute to confounding
  - A weight is assigned to each person according to the inverse of the conditional probability of his/her exposure status given the confounder(s)
Estimating the IPW: Example 1

Weight these people by \( \frac{1}{\text{pr}(A^- \mid L^+)} \)
Estimating the IPW: Example 2

Study Population (N)

L+

A+

Y+

Y-

A-

Y+

Y-

L-

A+

Y+

Y-

A-

Y+

Y-

Weight these people by $1 / \text{pr}(A+ | L-)$
**Estimating the IPW**

Example SAS code:

```sas
**estimate propensity scores**;
proc logistic descending data=dataset1;
   model A=L1 L2 L4 L5 L7 L8;
   output out=probA1_PS p=prA1_PS xbeta=logit_ps;
run;

data weights_PS;
   set probA1_PS;
   prA0_PS=1-prA1_PS;
   if A=1 then IPW=1/prA1_PS;
   else IPW=1/prA0_PS;
run;
```
Applying the IPW

Example SAS code:

**IP weighted RR in total population**;
proc genmod data=weights_PS;
   class ID;
   model Y=A /dist=poisson;
   weight IPW;
   repeated subject=id;
   estimate 'beta' A 1;
run;
How do these methods relate to RCTs?

- They use weighting to mimic randomization
- PS methods estimate effect using weighting across $x$ number of ‘new populations’
- IPW weights people by the IPW, creating a pseudo-population
The pseudo-population

Study Population (N)

L+

A+

Y+

Y-

A-

Y+

Y-

L-

A+

Y+

Y-

A-

Y+

Y-
The pseudo-population

Study Population (N)

- L+
  - A+
    - Y+
    - Y-
  - A-
    - Y+
    - Y-
- L-
  - A+
    - Y+
    - Y-
  - A-
    - Y+
    - Y-
Some practical considerations
Propensity score balancing
Propensity score balancing
How will PS be estimated?

For example:

- Covariates to be included
- Interaction terms to be included
- Type of model used to estimate the probability of exposure
- Type of regression used to model the outcome
How will IPW be estimated?

For example:

- Covariates to be included
- Interaction terms to be included
- Type of model used to estimate the probability of exposure
- Type of regression used to model the outcome
- Stabilized or unstabilized weights
- Robust variance estimation
To estimate $\text{RD}_{\text{causal}}$ in total population

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Traditional methods</th>
<th>Propensity score control</th>
<th>IPW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requires all confounding factors measured?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Statistical correctly specified</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Can handle large number of confounding variables?</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Requires homogeneity?</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Requires positivity?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
To estimate $RD_{\text{causal}}$ in treated

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Traditional methods</th>
<th>Propensity score control</th>
<th>IPW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requires all confounding factors measured?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Statistical correctly specified</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Can handle large number of confounding variables?</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Requires homogeneity?</td>
<td>Yes</td>
<td>No*</td>
<td>No#</td>
</tr>
<tr>
<td>Requires positivity?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

* If PS matching or SMR weighting to combine PS strata
# If SMR weighting used to create pseudo-population
Instrumental variable analysis: A solution?

- Stay tuned...
What if we are interested in a time varying treatment effect?

- Treatment (A) varies with time
What if we are interested in a time varying treatment effect?

- Treatment (A) varies with time

the causal effect if everyone in the population had been $A_{++}$ versus if everyone had been $A_{--}$
What if we are interested in a time varying treatment effect?

- Treatment (A) varies with time
- Covariate (L) varies with time
What if we are interested in a time varying treatment effect?

- Treatment (A) varies with time
- Covariate (L) varies with time and:
  - Predicts the outcome
What if we are interested in a time varying treatment effect?

- Treatment (A) varies with time
- Covariate (L) varies with time and:
  - Predicts the outcome
  - Predicts exposure
What if we are interested in a time varying treatment effect?

- Treatment (A) varies with time
- Covariate (L) varies with time and:
  - Predicts the outcome
  - Predicts exposure
  - Is affected by previous exposure

How do we address the confounding and mediational effects of L?
Inverse probability weighting: A solution?

- Stay tuned...
Take home messages

- These weighting approaches require all confounding factors are hypothesized and measured.
- To estimate the causal effect in the total population, only IPW does not require homogeneity of the treatment effect.
- The choice of confounding control method for point-treatment confounding should be driven by:
  - Causal question / target population
  - Practical considerations
Hernan MA and Robins JM. Estimating causal effects from epidemiological data. J Epidemiol Community Health 2006;60:578-86

Hernan MA et al. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology 2000;11:561-70


Sato T and Matsuyama Y. Marginal Structural Models as a Tool for Standardization. Epidemiology 2003;14:680-6

Sturmer T, Rothman KJ, Glynn RJ Insights into different results from different causal contrasts in the presence of effect-measure modification. Pharmacoepidemiol Drug Saf 2006; 15, 698-709
Thank you