Lessons about Confounding and Selection Bias Taught by the New User Design

Til Stürmer, FISPE
UNC Gillings School of Global Public Health
University of North Carolina at Chapel Hill, USA

Conflict of Interest

- I receive investigator initiated funding (R01 AG023178) from the National Institute on Aging of the NIH
- I receive research funding as PI of the UNC-DEcIDE Center of the US Agency for Health Care Research and Quality (AHRQ)
- I receive salary support as the director of the UNC Center of Excellence in Pharmacoepidemiology and Public Health
- I receive salary support from unrestricted research grants from pharmaceutical companies (Merck, Sanofi-Aventis) to UNC
- I do not accept personal compensation of any kind from any pharmaceutical company

Outline

- Why we need nonexperimental studies of medical interventions (brief)
- Confounding
  - By indication
  - By frailty
  - Study design options to limit confounding
  - Study design options to address unmeasured confounding
- Selection bias
  - By conditioning on persistence
  - Study design options to limit bias
Nonexperimental Studies of Medical Interventions
• Needed due to shortcomings of RCTs
  – Population, comorbidity, comedication, N, timeliness
  – E.g., benefit/harm of chemotherapy in > 75 year old
• Large population based linked healthcare databases
  – Solve all of the above
  – Sicker patients more likely to be treated

Confounding by Indication
• Good prescribing leads to confounding of drug effects on intended outcomes
• More severe disease more likely to
  – Be treated (with higher doses)
  – Have higher risk of adverse outcomes
• Assessment of severity of disease
  – Often difficult
  – Sometimes impossible
• Drug looks BAD compared with NON-USERS!

(Arche)typical Example for Confounding by Indication

Note: Problem is that we cannot measure asthma severity well (not the confounding per se)!
Confounding by Frailty in Population Based PE Studies

- Individuals close to death are
  - Less likely to receive preventive treatments
    - E.g., statins, flu vaccination
  - More likely to receive palliative treatments
    - E.g., opiates instead of NSAIDs
  - More likely to receive certain classes of drugs
    - E.g., loop diuretics vs. other diuretics
- Paradoxical drug mortality associations
- Drug looks GOOD vs. NON-USERS!

Frailty

- End of life loss of
  - Weight
  - Physical function
  - Cognitive function
- Recognized by healthcare professionals
- Reduces likelihood of preventive therapies
  - Focus on main medical problem (Redelmeier et al. 98)
  - Little expected benefit (competing risks; Welch et al. 96)

(Arche)typical Example for Confounding by Frailty

RR for 1-year mortality, hospitalized Medicare population.
(Arche)typical Example for Confounding by Frailty

Methods to Limit Confounding by Indication and Frailty
  • Compare initiators of treatments that are interchangeable within the same indication
  • Exclude patients without “equipoise” based on propensity scores/instrumental variables
  • Exchangeable cohorts of new users
  • Essentially eliminates confounding by frailty
  • Reduction of confounding by indication needs to be assessed for each comparator

Purely Hypothetical Studies
  • Insulin analogues and risk for cancer
    – Confounding by unmeasured BMI
  • Long-acting beta-agonists and risk for asthma mortality
    – Residual confounding by asthma severity
    – Archetype of confounding by indication
### Nonexperimental Study Comparing New Users of Insulin Analogue with Human Insulin

<table>
<thead>
<tr>
<th>BMI category</th>
<th>Insulin A</th>
<th>Insulin B</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;19</td>
<td>7 (0.9%)</td>
<td>14 (2.3%)</td>
</tr>
<tr>
<td>19–&lt;25</td>
<td>103 (13.1%)</td>
<td>115 (18.8%)</td>
</tr>
<tr>
<td>25–&lt;30</td>
<td>205 (26.0%)</td>
<td>151 (24.7%)</td>
</tr>
<tr>
<td>30–&lt;35</td>
<td>210 (26.6%)</td>
<td>156 (25.5%)</td>
</tr>
<tr>
<td>35–&lt;40</td>
<td>148 (18.8%)</td>
<td>84 (13.8%)</td>
</tr>
<tr>
<td>40–&lt;45</td>
<td>59 (7.5%)</td>
<td>48 (7.9%)</td>
</tr>
<tr>
<td>≥ 45</td>
<td>47 (6.0%)</td>
<td>39 (6.4%)</td>
</tr>
<tr>
<td>missing</td>
<td>10 (1.3%)</td>
<td>4 (0.7%)</td>
</tr>
</tbody>
</table>

*Note 1: this is NOT a randomized study!*
*Note 2: this will not ALWAYS work!*

### Nonexperimental Study Comparing New Users of LABA with ?

<table>
<thead>
<tr>
<th>Step</th>
<th>Option 1</th>
<th>Option 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low-dose ICS</td>
<td>Cromolyn, LTRA, Theophylline, Nedocromil, or Theophylline</td>
</tr>
<tr>
<td>2</td>
<td>Preferred: Low-dose ICS</td>
<td>Alternative: SABA PRN</td>
</tr>
<tr>
<td>3</td>
<td>Preferred: Low-dose ICS + LABA</td>
<td>OR Medium-dose ICS</td>
</tr>
<tr>
<td>4</td>
<td>Preferred: Medium-dose ICS + LABA</td>
<td>OR Low-dose ICS + LTRA, Theophylline, or Zileuton</td>
</tr>
<tr>
<td>5</td>
<td>Preferred: High-dose ICS + LABA</td>
<td>AND Consider Omalizumab for patients who have allergies</td>
</tr>
<tr>
<td>6</td>
<td>Preferred: High-dose ICS + LABA</td>
<td>AND Consider Omalizumab for patients who have allergies</td>
</tr>
</tbody>
</table>

### Limit Potential for Confounding by Study Design

- New user design (Kramer 87, Miettinen & Caro 89, Guess 89, Moride & Abenhaim 94, Ray 03)
- At a point in time where something happens
- Comparator drug (CER)
  - Therapeutic alternative
  - Same stage of disease progression
  - Same (similar) indication
- Confounding limited to choice between alternatives
- Ideally for same indication
  - Insulin analogue vs. NPH insulin: lack of control oral hypoglycemics
  - LABA vs. ? added to ICS: lack of control under ICS alone
  - Note: guidelines and expert knowledge needed here!
Additional Steps to Address Unmeasured Confounding

- Effect of unmeasured confounder on outcome of interest often well described
  - Reason to bother about BMI, asthma severity!
- Effect of unmeasured confounder on treatment decisions often unknown
  - Especially given measured covariates
  - May be estimated in external validation study using propensity scores (BMI data presented from EMR study)

What if BMI Not Balanced?

| Incident cancer burden attributable to excess body mass index in 30 European countries |
|---------------------------------|----------------------------------|-----------------|-----------------|
|                                | RR per 5kg/m²                   | 95% CI           | P value         |
| Risk factor                     | RR (95% CI)                     |                  |                 |
| Age                             | 1.34 (1.18, 1.52)               | < 0.001          |                 |
| Gender                          | 1.47 (1.30, 1.66)               | 0.001            |                 |
| Education                       | 1.19 (1.02, 1.39)               | 0.02             |                 |
| Smoking status                  | 1.87 (1.53, 2.28)               | < 0.001          |                 |
| Body mass index                 | 1.08 (1.05, 1.11)               | < 0.001          |                 |
| External Adjustment for Confounding

- Estimate effect of unmeasured confounder on treatment decision in (external) validation study
- Get estimate of the effect of the unmeasured confounder on the outcome from the literature
- "Adjust" main study estimate for unmeasured confounding using standard formulas (Bross 66)
- Separate estimates of the above for multiple confounders
- Weighted average of expected confounding
- Magnitude and overall direction of confounding
Adjusting for Unmeasured Confounders in Pharmacoepidemiologic Claims Data Using External Information

The Example of COX2 Inhibitors and Myocardial Infarction

<table>
<thead>
<tr>
<th>High School Education</th>
<th>Income (≥$50K)</th>
<th>Obese*</th>
<th>Age at Event</th>
<th>Current Users</th>
<th>Former Users</th>
<th>Never Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any COX2 inhibitor†</td>
<td>572</td>
<td>69</td>
<td>48</td>
<td>24</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>(Aspirin only)</td>
<td>562</td>
<td>69</td>
<td>46</td>
<td>24</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Referral only</td>
<td>244</td>
<td>66</td>
<td>48</td>
<td>19</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>Nonselective NSAIDs only</td>
<td>1382</td>
<td>72</td>
<td>56</td>
<td>24</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>251</td>
<td>73</td>
<td>58</td>
<td>23</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Naproxen</td>
<td>238</td>
<td>74</td>
<td>55</td>
<td>20</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Other NSAIDs only</td>
<td>1,077</td>
<td>71</td>
<td>55</td>
<td>25</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Nonusers</td>
<td>5,611</td>
<td>69</td>
<td>53</td>
<td>17</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

**Schneeweiss, Glynn, Tsai, Avorn, Solomon; Epidemiology 2005**

---

External Adjustment for Confounding

Practice of Epidemiology

Propensity Score Calibration in the Absence of Surrogacy

Mark Lent, Robert J. Glynn, Kenneth J. Rothman, Jerry Avorn, and Til Stürmer†

1 Correspondence to: Dr. Til Stürmer, Department of Epidemiology, UNC School of Public Health, University of North Carolina at Chapel Hill, 100 Manning Drive, CB 7505, Chapel Hill, NC 27599-7505, e-mail: stuermer@email.unc.edu

* Excluded are incomplete surveys, those with a household income less than $10,000, and those with age at event ≥60. The authors excluded these groups to improve the validity of the study results and to provide a more accurate representation of the population. The exclusion of these groups helps to control for potential confounding variables. The authors conclude that the results of the study are robust and can be generalized to the target population.

† The authors acknowledge the contributions of the following individuals: Dr. Robert J. Glynn, Dr. Kenneth J. Rothman, Dr. Jerry Avorn, and Dr. Til Stürmer. Their expertise and guidance were instrumental in the development of the study and the subsequent analysis. The authors express their sincere gratitude for the support and guidance provided.
Sensitivity Analysis for Unmeasured Confounding
Practice of Epidemiology

Selection Bias: Adherence to Placebo and Mortality

<table>
<thead>
<tr>
<th>Study</th>
<th>Good adherence to drug therapy</th>
<th>Poor adherence to drug therapy</th>
<th>Odds ratio (95% CI)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochrane Drug Policy Research Group (1999)</td>
<td>274/1813</td>
<td>246/682</td>
<td>27.02</td>
<td>7.02</td>
</tr>
<tr>
<td>Järvelin heart attack trial (1993)</td>
<td>311/137</td>
<td>447</td>
<td>10.1</td>
<td>0.35</td>
</tr>
<tr>
<td>Järvelin heart attack trial (1991)</td>
<td>114/220</td>
<td>421</td>
<td>4.29</td>
<td>0.27</td>
</tr>
<tr>
<td>Canadian arterial-renal vascular trial (1996)</td>
<td>42/104</td>
<td>1756</td>
<td>11.81</td>
<td>0.41</td>
</tr>
<tr>
<td>Scandinavian angina propranolol trial (1986)</td>
<td>8.88</td>
<td>140</td>
<td>1.54</td>
<td>0.26</td>
</tr>
<tr>
<td>Physicians health study (1980)</td>
<td>105/684</td>
<td>984/125</td>
<td>23.53</td>
<td>0.50</td>
</tr>
<tr>
<td>North of Scotland primary care study (1993)</td>
<td>80/330</td>
<td>65/373</td>
<td>15.56</td>
<td>0.67</td>
</tr>
<tr>
<td>University of Essen Diabetes Project (1992)</td>
<td>11/140</td>
<td>15/262</td>
<td>6.98</td>
<td>0.06</td>
</tr>
<tr>
<td>Total N (61)</td>
<td>154/429</td>
<td>629</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

OR for good vs. bad adherence and all cause mortality - Simpson, S. H et al. BMJ 2006

Selection Bias

- Better adherence associated with ~50% lower mortality (unless therapy harmful)
- Most likely explanation: sick stoppers
- Frail participants stop adhering to trial protocol
- Similar issue with persistence on preventive drugs?
Statin Persistence

Table 1. Association between Adherence to Statin Therapy and Risk of Health-Related Events

<table>
<thead>
<tr>
<th>Incident events</th>
<th>Non-adherent Rate</th>
<th>Less adherent Rate</th>
<th>Disaggregated Rate</th>
<th>95% Confidence Interval</th>
<th>Adjusted RR</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-statins-related events</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth</td>
<td>6.28</td>
<td>6.38</td>
<td>6.28</td>
<td>(0.11-0.67)</td>
<td>1.00</td>
<td>(0.79-1.26)</td>
</tr>
<tr>
<td>Fall</td>
<td>6.22</td>
<td>6.46</td>
<td>6.22</td>
<td>(0.09-0.36)</td>
<td>1.00</td>
<td>(0.68-1.31)</td>
</tr>
<tr>
<td>Fracture</td>
<td>2.28</td>
<td>2.36</td>
<td>2.28</td>
<td>(0.09-0.56)</td>
<td>1.00</td>
<td>(0.20-0.29)</td>
</tr>
<tr>
<td>Motor vehicle accident</td>
<td>1.00</td>
<td>2.20</td>
<td>1.00</td>
<td>(0.03-0.08)</td>
<td>1.00</td>
<td>(1.03-1.07)</td>
</tr>
<tr>
<td>Overdose</td>
<td>0.44</td>
<td>2.04</td>
<td>0.44</td>
<td>(0.09-0.09)</td>
<td>1.00</td>
<td>(0.95-1.05)</td>
</tr>
<tr>
<td>Rehospitalization</td>
<td>0.02</td>
<td>0.05</td>
<td>0.02</td>
<td>(0.01-0.03)</td>
<td>1.00</td>
<td>(0.99-1.01)</td>
</tr>
<tr>
<td>Medication-related events</td>
<td>1.21</td>
<td>2.32</td>
<td>1.21</td>
<td>(0.09-0.65)</td>
<td>1.00</td>
<td>(0.78-1.29)</td>
</tr>
<tr>
<td>All events</td>
<td>7.28</td>
<td>6.68</td>
<td>7.28</td>
<td>(0.10-0.54)</td>
<td>1.00</td>
<td>(0.83-1.25)</td>
</tr>
</tbody>
</table>

BC data; Dormuth et al. Circulation 2009

Healthy User Bias DAG

Healthy initiator → T₀ → T₁ → Y

Healthy adherer (sick stopper)

Conditioning on T₁ due to censoring if T₀=1 & T₁=0

Bias Over Time on Preventive Drug Compared with Non-Use

RR vs. no drug

1.0

Time on drug

Healthy initiator

Healthy adherer

Healthy user
Study Design Options to Limit Selection Bias

- Cannot be “controlled” for (introduced by control)!  
- Cannot be “avoided” by excluding stoppers-switchers-augmenters from cohort (introduced by)!  
- Solution: first treatment carried forward (ITT)  
  - Standard for RCT (efficacy): tends to bias towards null  
  - Works best for short term intended effects  
- Always present in as treated analyses, but:  
  - Limited by adding various lag times  
    - Reduces bias from informative censoring (extreme: ITT)  
  - Marginal structural models  
    - Dependent on prediction of treatment(s) based on all risk factors for the outcome

So How Does This All Fit Together?

IKEA Job Interview

Confounding and Selection Bias in Population Based Studies of Medical Interventions

- Study design  
  - New user design  
  - Comparator drug  
  - Sensitivity analyses on stopping/switching/augmenting  
  - Consider internal or external validation studies to control for unmeasured confounders  
  - Consider methods not based on no unmeasured confounding assumption (i.e., instrumental variables, self-controlled designs)  
- Study analysis  
  - Look for nonuniform effects e.g., over range of PS  
    - Consider PS matching, range restrictions, trimming  
    - Discuss unmeasured confounding vs. treatment heterogeneity
Take Home Message

• Confounding by indication and frailty major threats to validity of nonexperimental studies
• Selection bias major threat to validity of any study (not specific to nonexperimental studies)
• Advances in PE study design allow us to reduce potential for bias
  – New user design and active comparator cohorts
  – Implement various valid alternatives to handle changes in treatment during follow-up
• Proper research question and study design more important than analysis