The study of cardiovascular health outcomes in the era of claims data

Bruce M Psaty, MD, PhD
Introduction

• Event rates and data sources

• Cardiovascular Health Study
  – Events surveillance methods
  – Comparisons with claims data
  – Risk-factor associations
  – Example of an adverse drug reaction
  – Risk-factor predictions

• A few final observations
Variation in event rates

- The methods of case identification
- The intensity of the surveillance
- The aggressiveness of the data collection
- The criteria for adjudication
- The quality control of the overall effort
Administrative claims data

- Electronic side effects of billing and clinical care
- Economic incentives drive coding practices in claims – Irresponsible to ignore them, illegal to abuse them
- Useful as pointers to potential events
Cardiovascular Health Study

• Cohort study of 5888 older adults to evaluate risk factors for CHD and stroke

• Medicare files used to sample participants at four sites in 1989-90

• Broad interest in a variety of health conditions that affect older adults

CHS events data collection

• Self report of all events every 6 months supplemented by search of Medicare data

• Discharge summaries of all hospitalizations plus targeted data for specific events

• Broad set of codes plus Field-Center review to identify cases sent for full adjudication

• Each potential case of any event type was reviewed for all cardiac-event types

<table>
<thead>
<tr>
<th>Event</th>
<th>Number</th>
<th>Kappa</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI</td>
<td>241</td>
<td>0.86</td>
<td>0.78-0.93</td>
</tr>
<tr>
<td>Angina</td>
<td>241</td>
<td>0.79</td>
<td>0.71-0.86</td>
</tr>
<tr>
<td>HF</td>
<td>241</td>
<td>0.85</td>
<td>0.78-0.92</td>
</tr>
<tr>
<td>PAD</td>
<td>241</td>
<td>0.82</td>
<td>0.70-0.93</td>
</tr>
<tr>
<td>Stroke</td>
<td>62</td>
<td>0.87</td>
<td>0.75-0.99</td>
</tr>
<tr>
<td>TIA</td>
<td>62</td>
<td>0.87</td>
<td>0.66-1.00</td>
</tr>
</tbody>
</table>
Conventions for claims-data use

• Validation efforts, typically small, examine the positive predictive value

• Studies select a limited set of codes and often restrict to the primary diagnosis

• MI in any position, but only first position for stroke and HF codes
Primary codes for CMS non-events?

- For MI: HF, AF, hypertension, urinary tract infection, pneumonia, and COPD
- For HF: CHD, hypertension, AF, diabetes, COPD, pneumonia, anemia
- For stroke: hypertension, AF, CHD, HF, anemia, psychosis
Association & misclassification

• Low sensitivity reduces power
  – A loss that may not matter with big data

• The effects of low specificity depend on the risk factor’s association with the misclassified event: MI vs ACS

• CV risk factors tend to be promiscuous and themselves reasons for hospitalization
# Statin related rhabdomyolysis

<table>
<thead>
<tr>
<th>Data source</th>
<th>Simvastatin vs other statin</th>
<th>Simvastatin 80 vs 20 mg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validated cases</td>
<td>2.6</td>
<td>12.2</td>
</tr>
<tr>
<td>95% CI</td>
<td>(1.0-7.8)</td>
<td>(3.6-52)</td>
</tr>
<tr>
<td>Administrative data</td>
<td>1.0</td>
<td>1.8</td>
</tr>
<tr>
<td>95% CI</td>
<td>(0.8-1.3)</td>
<td>(1.1-2.9)</td>
</tr>
</tbody>
</table>

Floyd J. JAMA 2012; 307: 1580-2; PPV of 728.88 = 7.5%
Prediction & misclassification

• The effects of low specificity again depend on the risk factor’s association with the misclassified event
  – An effect that may weaken the prediction of genuine cardiovascular events

• Low sensitivity reduces event rates
  – A loss that matters for the estimation of the predicted absolute risk
Lipid guidelines controversy

• NHBLI gave up on ATP4 lipid guidelines
  – Long process that was transferred to the ACC/AHA and then rushed out last fall

• News headlines featured major concerns
  – Risk prediction tool “overestimated” risk
  – Threshold increased number of treated

• Data from CHS, FHS and ARIC were used to develop the risk prediction tool
REGARDS: missed events

• Events data, ascertained by self report rather than “active” surveillance, at first suggested “over” prediction of risk

• The investigators themselves had identified 345 events among those with Medicare data
  – Medicare data identified an additional 112 events, 24% of the total of 457 events

## Events rates in diabetes trials

<table>
<thead>
<tr>
<th>Study name</th>
<th>ARIC</th>
<th>RECORD</th>
<th>ACCORD</th>
<th>VADT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study type</td>
<td>cohort</td>
<td>trial</td>
<td>trial</td>
<td>trial</td>
</tr>
<tr>
<td>Age, years</td>
<td>59</td>
<td>58</td>
<td>62</td>
<td>60</td>
</tr>
<tr>
<td>Sex, % male</td>
<td>47</td>
<td>52</td>
<td>62</td>
<td>97</td>
</tr>
<tr>
<td>Diabetes, %</td>
<td>8</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Heart disease, %</td>
<td>0</td>
<td>32</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>Body mass index</td>
<td>27</td>
<td>31</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>Systolic pressure</td>
<td>118</td>
<td>139</td>
<td>137</td>
<td>132</td>
</tr>
<tr>
<td>LDL cholesterol</td>
<td>140</td>
<td>127</td>
<td>104</td>
<td>108</td>
</tr>
<tr>
<td><strong>Heart attack rate</strong></td>
<td><strong>4.3</strong></td>
<td><strong>4.6</strong></td>
<td><strong>13.8</strong></td>
<td><strong>15.9</strong></td>
</tr>
</tbody>
</table>

A few final observations

• Best not to assume that event rates are complete and accurate
  – Use events rates as a QC metric

• Event rates depend on intensity of identification and investigation

• Events data collection methods should be fit for purpose