Historical controls: think cluster not parallel

Stephen Senn, Olivier Collignon, Anna Schritz
Acknowledgements

Many thanks for the invitation

This work is partly supported by the European Union’s 7th Framework Programme for research, technological development and demonstration under grant agreement no. 602552. “IDEAL”
TARGET study

• Trial of more than 18,000 patients in osteoarthritis over one year or more
• Two sub-studies
  • Lumiracoxib v ibuprofen
  • Lumiracoxib v naproxen
• Stratified by aspirin use or not
• Has some features of a randomised trial but also some of a non-randomised study

<table>
<thead>
<tr>
<th>Demographic Characteristic</th>
<th>Sub-Study 1</th>
<th>Sub Study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of low-dose aspirin</td>
<td>975 (22.3)</td>
<td>1195 (25.1)</td>
</tr>
<tr>
<td>History of vascular disease</td>
<td>393 (9.0)</td>
<td>588 (12.4)</td>
</tr>
<tr>
<td>Cerebro-vascular disease</td>
<td>69 (1.6)</td>
<td>108 (2.3)</td>
</tr>
<tr>
<td>Dyslipidaemias</td>
<td>1030 (23.5)</td>
<td>799 (16.9)</td>
</tr>
<tr>
<td>Nitrate use</td>
<td>105 (2.4)</td>
<td>181 (3.8)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demographic Characteristic</th>
<th>Sub-study (DF=1)</th>
<th>Treatment given Sub-study (DF=2)</th>
<th>Treatment (DF=2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of low-dose aspirin</td>
<td>&lt; 0.0001</td>
<td>0.94</td>
<td>0.0012</td>
</tr>
<tr>
<td>History of vascular disease</td>
<td>&lt; 0.0001</td>
<td>0.07</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Cerebro-vascular disease</td>
<td>0.0002</td>
<td>0.93</td>
<td>0.0208</td>
</tr>
<tr>
<td>Dyslipidaemias</td>
<td>&lt;0.0001</td>
<td>0.92</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Nitrate use</td>
<td>&lt; 0.0001</td>
<td>0.10</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

(c) Stephen Senn
TARGET odds ratios CV event

Lumiracoxib v Lumiracoxib

<table>
<thead>
<tr>
<th>Outcome Variables</th>
<th>Deviance</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total of discontinuations</td>
<td>13.61</td>
<td>0.0002</td>
</tr>
<tr>
<td>CV events</td>
<td>2.92</td>
<td>0.09</td>
</tr>
<tr>
<td>At least one AE</td>
<td>1.73</td>
<td>0.19</td>
</tr>
<tr>
<td>Any GI</td>
<td>21.31</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Dyspepsia</td>
<td>47.34</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
A big data analyst is an expert at reaching misleading conclusions with huge data sets, whereas a statistician can do the same with small ones.
Data Filtering Some Examples

• Oscar winners lived longer than actors who didn’t win an Oscar
• A 20 year follow-up study of women in an English village found higher survival amongst smokers than non-smokers
• Transplant receivers on highest doses of cyclosporine had higher probability of graft rejection than on lower doses
• Left-handers observed to die younger on average than right-handers
• Obese infarct survivors have better prognosis than non-obese
Moral

• What you don’t see can be important
• For some purposes just piling on data does not really help
• What helps are
  • Careful design
  • Thinking!
• The TARGET study provides non-randomised control data that will be as goods as (in practice much better) than any historical data you will find
  • These data are still not good enough

(c) Stephen Senn
Cluster randomised

Randomisation

Centre

Centre

Centre

Centre

D  D  D  D  A  A  A  C  C  C  C  B  B  B  B
Moral

• We have a tendency to think that historical controls are slightly inferior concurrent controls
• Has led some to just propose a naïve discounting
• However such controls were not treated under similar conditions in the same centres
• They were treated in different centres
• At the very least we have the variability of a cluster-randomised trials
• In practice things will be worse
Implications for using historical controls

• Identification, pre-specification and agreement on a suitable historical data-set
  • Because otherwise you could pick and choose your historical controls

• An agreed, enforceable and checkable plan for recruiting the experimental arm in advance of doing so
  • Because otherwise you could selectively recruit to your advantage

• A finalised analysis plan prior to beginning the trial
  • Because blinding is impossible

• Use of a hierarchical model with sufficient complexity
  • Because many components of variation are involved

• Emphasis on number of historical trials rather than patients
  • Because otherwise components of variation cannot be estimated
We tend to believe “the truth is in there”, but sometimes it isn’t and the danger is we will find it anyway
Suggested reading


Galwey, N.W., 2016. Supplementation of a clinical trial by historical control data: is the prospect of dynamic borrowing an illusion?. *Statistics in Medicine*.
