Biostatistics, Medical Research, and Medical Ethics

Timothy E. O’Brien, Ph.D.
Department of Mathematics & Statistics
Loyola University Chicago

AEA Talk 2nd May 2006

Acknowledgements –
Jerry Avorn, MD, Harvard Medical School,
“Powerful Medicines”,
and
Marcia Angell, MD, Harvard School of Social Med.,
“The Truth About the Drug Companies”
Talk Outline

I. Basic Biostatistical Methods –
   A. Objectives and Uses
   B. ANOVA
   C. Regression
   D. Analysis of CDA data

II. Medical and Pharmaceutical Research
   A. Phases of Drug Discovery and Testing
   B. A Quick History
   C. Assessing (Actual) Efficacy
   D. Assessing Risk and Safety
   E. Assessing Costs

III. Some Ethical Considerations

IV. Conclusions
I. Basic Biostatistical Methods

A. Objectives and Uses

- Modelling
- Estimation
- (Hypothesis) Testing

B. ANOVA

1. 60 pts. and 2 anti-hypertension drugs (A & B), then measure Y = SBP after 24 hours. Could randomize 30 pts. to A and 30 pts. to B. Assume (1) Gaussian dist., (2) equal variances, (3) independent measurements, then do an *independent sample t-test*. 
2. 60 pts. and 3 anti-hypertension drugs (A, B & C), then measure Y = SBP after 24 hours. Could randomize 20 pts. to each drug. Assume (1) Gaussian dist., (2) equal variances, (3) independent measurements, then do a one-way ANOVA.

3. Back to 60 pts. and 2 anti-hypertension drugs (A & B) – but now with 30 sets of twins. A set of twins constitutes a “block” since the “Experimental Units” (EU’s) within a block are typically much more similar than those from two different sets of twins. Other examples of blocking include plots of land in a geographic region, litters of mice, etc. Then, one EU within each block is randomized to each of the treatments. In this instance, we can perform a paired t-test on the differences of the SBP within the blocks (pairs of twins), assuming Gaussian distribution.

4. Other ANOVA’s includes the analysis of repeated measures, crossover designs, etc.
C. Linear Regression

Simple Linear Regression – For example,

\[ SBP = \alpha + \beta \cdot CHOL + \varepsilon \]

Need to assume: (a) line is correct; (b) Gaussian distribution; (c) constant variance; (d) independent measurements.

Multiple Linear Regression – Could include (on the RHS) other potentially useful explanatory variable(s) in the model. For example,

\[ SBP = \alpha + \beta \cdot CHOL + \gamma \cdot AGE + \varepsilon \]
D. Analysis of Categorical Data

First example:

<table>
<thead>
<tr>
<th>Gender</th>
<th>Infected?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>Yes</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>12</td>
</tr>
<tr>
<td>Female</td>
<td>Yes</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>20</td>
</tr>
</tbody>
</table>

\[ \pi_M = \text{Probability of Infection for Males; } p_M = 0.52 \]
\[ \pi_F = \text{Probability of Infection for Females; } p_M = 0.20 \]
- can test \( \pi_M = \pi_F \) using a \( \chi^2 \) test
- can estimate \( RR = \frac{\pi_M}{\pi_F} \) and test \( RR = 1 \)
- can estimate \( OR = \frac{\pi_M/(1-\pi_M)}{\pi_F/(1-\pi_F)} \) and test \( OR = 1 \)

Second example (popular with epidemiological studies):

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Neural Tube Defects</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case</td>
<td>Control</td>
</tr>
<tr>
<td>TT</td>
<td>73</td>
<td>83</td>
</tr>
<tr>
<td>CC</td>
<td>151</td>
<td>439</td>
</tr>
</tbody>
</table>

\[ \pi_{TT} = \text{Probability of Case for TT; } p_{TT} = 0.4679 \]
\[ \pi_{CC} = \text{Probability of Case for CC; } p_{CC} = 0.2559 \]

\( \Rightarrow \) 95% CI for OR is (1.78, 3.68)
Caveat – these methods lack power for tables like:

<table>
<thead>
<tr>
<th>Pain Relief</th>
<th>Drug A</th>
<th>Drug B</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Some</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Substantial</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Complete</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>25</td>
</tr>
</tbody>
</table>

The (incorrect) commonly-used $\chi^2$ statistic here gives a p-value of 14.2% and the FET p-value is 16.2%. On the other hand, the (correct) Mantel-Haenszel $\chi^2$ p-value is 2.8% – thereby indicating the superiority of Drug A over B.

**Biostatistician’s Challenge** – beyond the challenge of modeling, is to match the correct “statistical tool” to the problem at hand, so as to answer the relevant question(s) via estimation and/or testing.
II. Medical and Pharmaceutical Research

A. Phases of Drug Discovery and Testing
   - Non-clinical (compound(s) in a Petri dish)
   - Pre-clinical (studies in rats)
   - Clinical
     - Phase I (PK, healthy volunteers, how drug metabolizes)
     - Phase II (dosing, safety, MTD)
     - Phase III (large-scale clinical trial for efficacy)
     - Phase IV (post-marketing, other indications)

   (often now, drug companies “buy” promising compounds from biotech companies, and then just do the clinical testing)

B. A Quick History (Avorn, Ch. 2)
   - Old days: potions; aspirin from willow bark, etc.
   - Giving government clout: FDA since 1962
   - Scientific Evolution: shift from anecdotal evidence to Randomized controlled trials (RCT’s); important to show an improvement over the placebo effect:
     
     My doctor gave me these pills and my symptoms were gone in a week! A good thing – otherwise the problem would have taken seven days to resolve.

   - Placebo effects are observed through MRI’s
   - Biostatistical revolution and p-values.

Avorn: each drug represents a triangle with 3 faces representing healing, hazards, and costs.
C. Assessing *(Actual)* Efficacy

- Is challenging because drug companies/FDA continued to resist head-to-head drug comparisons with other similar drugs.
- An inspiring exception: the ALLHAT study (see below).

*The net effect, though, is that drugs have been tested and approved based on relatively short randomized controlled clinical trials; later, numerous serious adverse experiences have surfaced.*

D. Assessing Risk and Safety

- “All medicines are poisons .. the right dose differentiates a poison from a remedy.” Avorn, p.72 (Med. School)
- “Every drug has at least two effects: the one you intended and the one you didn’t.”
- Once drugs are approved, cannot rely on (profit-focused) drug companies nor the (overwhelmed) FDA to monitor AE’s (adverse experiences, or side-effects).
- Once drugs have been approved and marketed to the public, risks can only be assessed using epidemiological tests such as *case-control studies*, and these *(observational studies)* are often marred by confounding factors and “confounding by indication.”
• Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) – head-to-head comparison (RCT) of anti-hypertensive drugs – showed the less-expensive diuretics safest and equally efficacious. This study contradicted the indications from previous case-control studies!

E. Assessing Costs
• very challenging: is illness in a 92-year-old “worth less” than the same illness in a 29-year-old? Is ED “worth the same” as a heart attack? What fraction? Who decides? How do we discount to present value a heart attack in 20 years?
• very important since the financial strain on State and Federal Medicare and Medicare programs is very great.
• felt most dearly by the elderly, who are often forced to travel to foreign lands or the Internet to fill drug prescriptions.
• Some hope – Avorn’s anecdote: “Academic detailing” (NEJM, 1983) under the auspices of the U.S. government.
III. Ethics (Drug Testing in the Third World)

• “The Constant Gardener” and “The Body Hunters” focus on drug testing in Africa and Asia
• The former is based on (true story): Pfizer tested its antibiotic Trovan in Kano, Nigeria, which was subsequently withdrawn: “after less than two years on the market, there were over a hundred reports that the drug produced liver toxicity, causing several deaths, and it is no longer available.”
• The latter story was the Washington Post 6-part series on drug testing in Africa and ethical concerns: http://www.washingtonpost.com/wp-dyn/world/issues/bodyhunters/
• Movie and article (p.2) address:
  - how drug companies distort research to make their drugs look safer and more effective than they are,
  - how they can get away with this more easily in poor regions of the world,
  - and how they use their vast wealth to influence governments and the medical profession and any other institutions that might interfere with their single-minded pursuit of profits.
• Angell: “.. it was unethical to test an experimental drug orally in the midst of an epidemic.” These tests and others conducted in the third world are “inherently exploitative.”
In the U.S.,
- NDA’s must be filed with the FDA
- IRB’s (Institutional Review Boards) must be established to monitor safety and side-effects
- IC (informed consent) must be obtained and continuously updated (IC is a dialogue rather than a form to sign)

No IRB was set up for the Pfizer study and IC was dubious at best.

AZT versus placebo studies in Africa and Thailand to test HIV transmission from mother to child!

IV. Some Conclusions and Hope:

Given our capitalistic system, can we really blame drug companies? Avorn lays a good deal of the blame on MD’s who compromise their striving to provide better health-care for financial gain. Equally culpable: politicians (even within FDA) for the same reason.

All is not hopeless given Avorn’s successes and our dire need!