How to Make Sense of Statistics
Reported in the Medical Literature

Colin P. West, MD, PhD
Division of General Internal Medicine
Division of Biomedical Statistics and Informatics

Denise M. Dupras, MD, PhD
Division of Primary Care Internal Medicine
Disclosures

• Dr. West has no known conflicts of interest to disclose.

• Dr. Dupras has no known conflicts of interest to disclose.
Session Objectives

• Highlight several common issues in published research that lead to misinterpretation of study findings.

• Promote the EBM principle of “enlightened skepticism” when reading medical literature.

• Challenge the dogma that EBM and medical statistics are dull by having fun and telling several bad jokes.
Tip #1: Clinical vs Statistical Significance

- **Background**
  - Statistical significance
    - UNLIKELY to occur by chance
    - “Is this a real effect?”
    - Not = important or meaningful
  - Clinical significance
    - “Is this effect important or meaningful?”
Tip #1: Clinical vs Statistical Significance

• Outcomes
  • If you can measure it, you can ask if a difference is statistically significant
  • Specific outcomes impact clinical significance
    • Change in systolic BP versus stroke
    • Visual analog scale for pain
Tip #1: Clinical vs Statistical Significance

• Example
  • Effect of statins on carotid intimal thickness
  • Meta-analysis – 10 studies, 3443 subjects
  • Results
    Beneficial effect of statins (p<0.00001)
    -0.02235 (-0.02656, -0.01614) mm/y

Tip #1: Clinical vs Statistical Significance

- Questions to ask yourself:
  - Would the reported effect matter to an individual patient?
  - Would the reported effect change public health if it were applied to a larger group of patients?
Tip #2: Absolute vs Relative Risk

- **Background**
  - Differences between 2 groups
  - Absolute = actual difference
- **Example**
  - Treatment group – 10%
  - Placebo group – 30%
  - Absolute risk reduction = 20%
- Number needed to Treat = 1/ARR
  - NNT = 5
Tip #2: Absolute vs Relative Risk

• Background
  • Relative = change in one group compared to another group (↑↓)

• Example
  • Treatment group – 10%
  • Placebo group – 30%
  • Relative risk = 1/3 or 33%
  • Relative risk reduction = 67%
Tip #2: Absolute vs Relative Risk

An Example

<table>
<thead>
<tr>
<th></th>
<th>Drug</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>No Event</td>
<td>90</td>
<td>50</td>
</tr>
<tr>
<td>% Event</td>
<td>10%</td>
<td>50%</td>
</tr>
<tr>
<td>ARR</td>
<td></td>
<td>40%</td>
</tr>
<tr>
<td>NNT</td>
<td>1/0.4</td>
<td>2.5</td>
</tr>
<tr>
<td>RRR</td>
<td>40%/50%</td>
<td>80%</td>
</tr>
</tbody>
</table>
Tip #2: Absolute vs Relative Risk

An Example

<table>
<thead>
<tr>
<th></th>
<th>Drug</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>No Event</td>
<td>90</td>
<td>50</td>
</tr>
<tr>
<td>% Event</td>
<td>10%</td>
<td>50%</td>
</tr>
<tr>
<td>ARR</td>
<td></td>
<td>40%</td>
</tr>
<tr>
<td>NNT</td>
<td>1/0.4</td>
<td>2.5</td>
</tr>
<tr>
<td>RRR</td>
<td>40%/50%</td>
<td>80%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Drug</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>No Event</td>
<td>9990</td>
<td>9950</td>
</tr>
<tr>
<td>% Event</td>
<td>0.1%</td>
<td>0.5%</td>
</tr>
<tr>
<td>ARR</td>
<td></td>
<td>0.4%</td>
</tr>
<tr>
<td>NNT</td>
<td>1/0.004</td>
<td>250</td>
</tr>
<tr>
<td>RRR</td>
<td>0.4%/0.5%</td>
<td>80%</td>
</tr>
</tbody>
</table>
Tip #2: Absolute vs Relative Risk

• What do pharmaceutical ads report?
• What is reported on the news?
Tip #3: P Values and Confidence Intervals

• Background
  • What is a p value anyway?
    • The probability of rejecting the null hypothesis (typically that two treatments do not differ) IF IN FACT THEY DO NOT DIFFER.
    • It is NOT the probability that the study finding is false!
Tip #3: P Values and Confidence Intervals

• **Background**
  
  • What does this mean?
  
  • If we think the treatment effect is implausible going into the study, even a significant p value may not make the effect likely.
  
  • This is analogous to interpretation of a positive diagnostic test in a patient for whom the diagnosis is very unlikely – the test often does not confirm the diagnosis!
  
Tip #3: P Values and Confidence Intervals

• **Confidence interval**
  • A measure of the precision of your estimates.
  • A 95% confidence interval for your outcome means (roughly) that 95% of similarly conducted studies would yield an interval that includes the true estimate.
  • Often interpreted as “the true effect is in this interval, with 95% confidence”
  • Not quite right, but close enough for many circumstances.
Tip #3: P Values and Confidence Intervals

• One useful tip is to look at the confidence interval associated with the reported p value
  • If very wide, the results may not be as trustworthy.
  • May be due to small sample size or small number of outcomes
    • In either case, USE CAUTION!
  • May be due to implausible or improbable association
    • If it seems too good to be true …
Tip #3: P Values and Confidence Intervals

• Another useful tip is to consider the clinical impact of each end of the confidence interval
  • If these are similar, the confidence interval is precise enough
  • If these differ markedly, however, the confidence interval lacks precision sufficient for clear clinical decision-making
Tip #3: P Values and Confidence Intervals

- Consider the precision of each of the following 95% CI’s for a relative risk, for which “no effect” is indicated by RR=1:
  - (11,15)
  - (1.1,212)
  - (0.1,14)
  - (0.97,1.02)
Tip #3: P Values and Confidence Intervals

• Example:
  • PPA and risk of hemorrhagic stroke
    • In women taking high doses as appetite suppressant (not as a cold medication), OR 16.58, p value 0.02.
    • 95% CI: (1.51, 182.21)!

Phenylpropanolamine and the risk of hemorrhagic stroke.

Tip #4: Absence of Evidence Is Not Evidence of Absence

• Another Example:
  • Based on a nonsignificant statistical test it may be declared that there is no association between an exposure and an outcome.
  • Are these confidence intervals for a RR telling us the same thing?
    • (0.99, 1.01)
    • (0.50, 2.75)
Tip #4: Absence of Evidence Is Not Evidence of Absence

• Example:
  • Pelvic lymphadenectomy vs. none in early-stage endometrial carcinoma
    • RR of death was 1.20 (95% CI 0.70, 2.07), p=0.50.
    • Reported as “similar risks”
      • C/w 30% risk reduction or doubling of risk, both clinically important

Tip #5: Multiple Testing and the Isolated Significant P Value

• **Background**
  - Assuming the null hypothesis is true, the probability of falsely rejecting it with a p value of 0.05 is 5%, and the probability of correctly rejecting it is 95%.
  - If we run 20 tests, the probability all of them are “correct” is \((0.95)^{20} = 0.36\).
  - So there is a 64% chance at least one of the conclusions when 20 tests are run is WRONG – and we can’t tell which one it is!
Tip #5: Multiple Testing and the Isolated Significant P Value

- The tip is to look skeptically at a table of many results when one is selected out of many for further discussion.
  - Suspect data dredging, or “torturing the data until it confesses”
  - What makes this more believable?
    - Biologically plausible association
    - Prior evidence
    - Clear a priori primary hypothesis
Tip #5: Multiple Testing and the Isolated Significant P Value

• Example:
  • Caffeine and the risk of breast cancer
    • In women with benign breast disease consuming at least 4 cups of coffee daily, RR for breast cancer was 1.35 (95% CI 1.01-1.80).
  • 230 statistical tests reported
    • 18 statistically significant results
    • Would expect 12 by chance alone, and very few of the results were “strongly significant”

Summary

- 5 Tips for “enlightened skepticism”
  - Clinical vs statistical significance
  - Absolute vs relative risk
  - P values and confidence intervals
  - Absence of evidence is not evidence of absence
  - Multiple testing and the isolated significant p value

- Thank You!