EBM Basics
EBM Vs Research Literacy
Meta-Analyses

Systematic Reviews

Randomized Controlled Clinical Trials

Randomized Experimental Studies

Case Control, Cohort, Cross Sectional Quasi-Experimental

Case Studies, Basic Research, Observational (Field Studies)
Why teach research literacy?

- Survey of IM residents with a multiple choice test on biostatistics and study design
- N=277
- “Most residents in this study lacked the knowledge in biostatistics needed to interpret many of the results in published clinical research.”
Table 3. Percentages of Correct Answers for the Knowledge-Based Questions

<table>
<thead>
<tr>
<th>Objective Correct (95% CI), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a. Identify continuous variable</td>
</tr>
<tr>
<td>1b. Identify ordinal variable</td>
</tr>
<tr>
<td>1c. Identify nominal variable</td>
</tr>
<tr>
<td>2. Recognize a case-control study</td>
</tr>
<tr>
<td>3. Recognize purpose of double-blind studies</td>
</tr>
<tr>
<td>4a. Identify ANOVA</td>
</tr>
<tr>
<td>4b. Identify 2 analysis</td>
</tr>
<tr>
<td>4c. Identify t test</td>
</tr>
<tr>
<td>5. Recognize definition of bias</td>
</tr>
<tr>
<td>6. Interpret the meaning of P value .05</td>
</tr>
<tr>
<td>7. Identify Cox proportional hazard regression</td>
</tr>
<tr>
<td>8. Interpret standard deviation</td>
</tr>
<tr>
<td>9. <strong>Interpret 95% CI and statistical significance</strong></td>
</tr>
<tr>
<td>10. Recognize power, sample size, and significance-level</td>
</tr>
<tr>
<td>Relationship</td>
</tr>
<tr>
<td>11. Determine which test has more specificity</td>
</tr>
<tr>
<td>12. Interpret an unadjusted odds ratio</td>
</tr>
<tr>
<td>13. Interpret odds ratio in multivariate regression analysis</td>
</tr>
<tr>
<td>14. Interpret relative risk</td>
</tr>
<tr>
<td>15. Determine strength of evidence for risk factors</td>
</tr>
<tr>
<td>16. Interpret Kaplan-Meier analysis results</td>
</tr>
</tbody>
</table>
What happens if we don’t?

- CV4 contraindicated or relatively contraindicated in pregnancy
- Based on an abstract published in 1992 on 6 post-dates women, showing onset of uterine contractions after a single CV4 (1.5-34 min)
- JAOA Vol 92, No 9.
Objectives

- Begin to critically read research
- Become familiar with the language of research
- Recognize how research relates to EBM
- Recognizing study design
- Recognizing statistical tests used
- Reading and interpreting statistical output and findings
- Knowing where in an article to find this information
How do you start making sense of a research article?

- Identify concepts in the article
- Use context clues to help understand the type of study done
- Understand strengths and weakness of the study type
- Know where it fits in the hierarchy of evidence
- Know the interpretation of the findings
- Identify what evidence was presented in the findings in support or opposition of the hypothesis
So what do we need to know?

- Research design
  - Observational
  - Case Control
  - Cohort
  - Cross sectional
  - Experimental
- Power
- Effect size
- Significance
  - P value
  - Confidence interval
- Epidemiology
  - Prevalence
  - Incidence
- Types of variables
  - Nominal
  - Ordinal
  - Continuous
- Common statistics
  - ANOVA
  - Chi-square
  - T-test
- Bias
- Validity
- Reliability
- Confounding
Example A: Study Design

- In the Licciardone JC, et al. study “Osteopathic manipulative treatment of back pain and related symptoms during pregnancy: A randomized controlled trial,” which of the following is a feature of the study design?

A. The subjects choose the treatment group to participate in

B. The subjects were assigned to treatment groups by the investigators

C. Subjects with back pain were identified and compared to subjects with no back pain
Example B: Study Design

To study the use of OMM for low back pain in pregnancy, you have identified a group of pregnant women in the last month of pregnancy. You are going to give them a survey and ask if they had back pain while pregnant, what types of treatments they tried, and if those treatments helped their pain. What type of study could this be?

A. Randomized Controlled Trial
B. Cohort Study
C. Cross-Sectional
TABLE
Baseline characteristics of randomly assigned subjects according to treatment group

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Treatment group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UOBC + OMT (n = 49)</td>
</tr>
<tr>
<td>Age, y</td>
<td>23.8 ± 5.5</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>23 (47)</td>
</tr>
<tr>
<td>Black</td>
<td>10 (20)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>15 (31)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Education, y</td>
<td>12.1 ± 1.7</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>29 (59)</td>
</tr>
<tr>
<td>Married</td>
<td>17 (35)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (6)</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>20 (41)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>24 (49)</td>
</tr>
<tr>
<td>Status unknown</td>
<td>5 (10)</td>
</tr>
<tr>
<td>Health insurance type</td>
<td></td>
</tr>
<tr>
<td>Medicaid</td>
<td>31 (63)</td>
</tr>
<tr>
<td>HMO/PPO/POS</td>
<td>14 (29)</td>
</tr>
<tr>
<td>Other</td>
<td>4 (8)</td>
</tr>
</tbody>
</table>

Types of Observational Studies

**Study Type**

- **Case Control**
  - Identify Subjects with Condition and No Condition
  - For each subject, determine if they had the risk factor or did not have the risk factor

- **Cohort**
  - Identify Condition Free Subjects with Risk Factor or No Risk Factor
  - For each subject, determine if they develop the condition or do not develop the condition

- **Cross Sectional**
  - Identify Subjects
  - For each subject determine if they had the risk factor and if they had the condition
Case-control study asks
- What explains why one group developed a condition and another did not
- Examines two groups (one with and one without specified conditions)

Cohort studies a group over time prospectively and may ask how that group develops a condition

Cross Sectional is a study of a population at a point in time
Subjects identified as having condition

Subjects identified as not having condition

Presence of Risk Factor Determined

Case Control Study
Condition free subjects identified as having risk factor

Condition free subjects identified as not having risk factor

Development of Condition Determined

Cohort Study
Subjects identified

Cross Sectional Study

Presence of Condition and Risk Factor Determined

No Condition Condition Risk Factor No Risk Factor
Strengths and Weaknesses

- **Case Control**
  - Good for rare conditions
  - Usually quick results, no follow up, inexpensive
  - Cases in study may not represent all cases
  - Bias in measuring risk factors after condition determined

- **Cohort**
  - Good for rare risk/exposure and evaluating risk prior to disease
  - Can study multiple outcomes at once
  - Can learn about incidence
  - May need many subjects or have long follow up period

- **Cross Sectional**
  - Generally one time measurement, quick results, inexpensive
  - Can establish prevalence, but can not address incidence
  - Cannot decide if risk preceded development of condition, or other way around
Prevalence and Incidence

- **Prevalence**
  - Proportion of individuals who have the condition at a given point in time
  - Provides an estimate of risk that an individual will be ill at a point in time
    - \( P = \# \text{ existing cases at a point in time} \)
    - \( P = \# \text{ of existing cases at a point in time} \)
    - \( P = \# \text{ existing cases at a point in time} \) population at that same point in time
  - Dependent on incidence and duration

- **Incidence**
  - Number of new events in a population during a time period
    - \( I = \# \text{ of new cases during the time period} \)
    - \( I = \# \text{ of new cases during the time period} \) population at risk
Prevalence and Incidence

- **Interaction:**
  - If incidence is low but duration is long, prevalence is relatively high
  - If incidence is high but duration is short, prevalence is relatively low

- **Impact of medical advances and discovery:**
  - Prevention: Incidence decreases, prevalence decreases
  - Treatment: Incidence same, prevalence decreases
  - Survival: Incidence same, prevalence increases
Power

- Protects against Type II error
- Type II error
  - wrongly concluding something is not there when it is
  - the researchers missed finding a difference or relationship that exists
- Power
  - chance of finding a meaningful difference if one really exists
  - defined as 1-probability of a Type II error
- There is a minimum power that most researchers will accept
Reading the article: Power

- Power is related to study design (not data analysis)
- Look in the methods section to see if Power is discussed and reported
- Look in the methods to see if a “Power Analysis” or “Sample Size Calculation” was performed
- The larger the Power the less likely the chance of making a Type II error
- A general rule: Power greater than 80% preferred
Effect Size

- What is a meaningful effect size?
  - How strong does a relationship have to be for us to consider it meaningful?
  - How big does a difference have to be for us to consider it meaningful?

- Meaningful Effect Size: difference between groups or the strength of a relationship the investigator deems clinically important

- There is a minimum effect size that most researchers will accept

- The larger an effect size the easier it is to observe it

- Effect size expected in a study is based on previous experience, previous studies, and clinical judgment
Reading the article: Effect Size

• Look in the methods section to see if meaningful effect size was discussed prior to doing the study and how it was determined

• Look in the results and findings to see if effect size for this study was reported

• The concept of meaningful effect size is used often in clinical practice
  - “You should lower your cholesterol by 10%”
  - “Quitting smoking will increase survival by 5 years”
Significance and $P$

- Protects against Type I error
- Type I error
  - wrongly concluding something is true when it is just random chance
  - the researchers think they have found something, but it isn’t really there
- There is a maximum chance of Type I error that the researcher will accept
\( \alpha \) and \( P \) values

- \( \alpha \) (alpha)
  - the maximum Type I error that is acceptable
  - Determined ahead of time
  - used to evaluate if \( P \) value is “statistically significant”

- \( P \) value
  - estimate of how likely it is the results are due to random chance rather than meaningful relationship
  - If \( P \) value is less than \( \alpha \), the results are said to be “statistically significant”
Small $P$ values for findings indicate only a small probability that results are due to random chance.

A general rule: $P$ values less than .05 are considered significant.

In tables or charts, significant $P$ values may be identified with notations.

Look in the methods section to see if the authors specified what $P$ value is considered significant or state what $\alpha$ is for the study.

### Table 3

Chronic Obstructive Pulmonary Disease: Mean (SD) Percent Predicted Value of Pulmonary Function Parameters for the OMT and Sham Therapy Groups (N=35)

<table>
<thead>
<tr>
<th>Pulmonary Function Parameter</th>
<th>OMT Group (n=18)</th>
<th>Sham Group (n=17)</th>
<th>P Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretreatment</td>
<td>Posttreatment</td>
<td>Pretreatment</td>
</tr>
<tr>
<td>FEV₁</td>
<td>45 (23)</td>
<td>44 (22)</td>
<td>46 (20)</td>
</tr>
<tr>
<td>FVC</td>
<td>73 (23)</td>
<td>69 (23)</td>
<td>77 (20)</td>
</tr>
<tr>
<td>FEF₅₀%</td>
<td>26 (18)</td>
<td>23 (15)</td>
<td>26 (19)</td>
</tr>
<tr>
<td>FEF₂₅₇₅%</td>
<td>17 (12)</td>
<td>16 (11)</td>
<td>16 (11)</td>
</tr>
<tr>
<td>FEF₇₅%</td>
<td>24 (21)</td>
<td>17 (12)</td>
<td>20 (13)</td>
</tr>
<tr>
<td>FEF₂₅₇₅%</td>
<td>18 (11)</td>
<td>16 (10)</td>
<td>18 (12)</td>
</tr>
<tr>
<td>FEF₃₅₅₉₃₅%</td>
<td>53 (25)</td>
<td>57 (23)</td>
<td>68 (30)</td>
</tr>
<tr>
<td>FIF₅₀%</td>
<td>63 (35)</td>
<td>63 (24)</td>
<td>74 (32)</td>
</tr>
<tr>
<td>ERV</td>
<td>85 (49)</td>
<td>73 (24)</td>
<td>70 (36)</td>
</tr>
<tr>
<td>IC</td>
<td>58 (18)</td>
<td>59 (19)</td>
<td>66 (18)</td>
</tr>
<tr>
<td>MVV</td>
<td>47 (24)</td>
<td>42 (22)</td>
<td>56 (17)</td>
</tr>
<tr>
<td>SVC</td>
<td>65 (18)</td>
<td>61 (18)</td>
<td>67 (16)</td>
</tr>
<tr>
<td>TGV</td>
<td>158 (57)</td>
<td>172 (87)</td>
<td>168 (44)</td>
</tr>
<tr>
<td>RV</td>
<td>190 (88)</td>
<td>218 (127)</td>
<td>212 (67)</td>
</tr>
<tr>
<td>TLC</td>
<td>114 (28)</td>
<td>124 (46)</td>
<td>122 (22)</td>
</tr>
<tr>
<td>Airway resistance</td>
<td>394 (365)</td>
<td>357 (275)</td>
<td>361 (215)</td>
</tr>
<tr>
<td>Airway conductance</td>
<td>24 (15)</td>
<td>26 (23)</td>
<td>26 (18)</td>
</tr>
</tbody>
</table>

* Statistical significance was tested using the nonparametric analysis of covariance. The pretreatment value was the covariate.
† Statistically significant difference in the percent predicted values for the study groups between pre- and posttreatment (P<.05).
Significance and Confidence Intervals (CI)

- Helps determine to what extent the findings support or refute the hypothesis
- Gives more information than just “significant” or “not significant”
- Based on sample size and the level of confidence the researchers consider acceptable
- In general:
  - Studies with larger sample sizes have narrower CI than those with small sample sizes
  - 90% CI are narrower than 95% CI (and so on)
  - Sometimes a value included in a CI is clinically meaningful, suggesting need for more study
Confidence Intervals

- 95% confidence interval – If the experiment were conducted with many different samples drawn the same way from the same population, 95% of these would include the population parameter.
  - We sometimes say we are 95% confident that this CI includes the true population value.
  - It does not mean that there is a 95% chance that the “truth” is in the interval - because it either is or it isn’t.

A 95% CI

Misses the true mean!

95% of our CI’s successfully covered the true mean!
General rule for using CI to assess significance:

- For a DIFFERENCE between groups: a CI that includes 0 indicates that the difference between groups is not statistically significant at that confidence level.
- For a RATIO: a CI that includes 1 indicates that the ratio is not statistically significant at that confidence level.

If both $P$ value and CI are reported in an article for an analysis – they should both support the same conclusion.
Table 4: Crude odds ratios (ORs) and 95% confidence intervals (CIs) for the associations between type 2 diabetes mellitus (T2DM) and osteopathic palpatory findings according to element of somatic dysfunction, spinal segmental level, and laterality.\(^\text{a}\)

<table>
<thead>
<tr>
<th>Element of Somatic Dysfunction</th>
<th>Spinal Segmental Level</th>
<th>Left</th>
<th>Right</th>
<th>OR (95% CI)</th>
<th>T2DM+</th>
<th>T2DM-</th>
<th>OR (95% CI)</th>
<th>T2DM+</th>
<th>T2DM-</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5–T7</td>
<td>16/44</td>
<td>7/25</td>
<td>1.30</td>
<td>(0.47–3.58)</td>
<td>15/45</td>
<td>7/25</td>
<td>1.19</td>
<td>(0.43–3.31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T8–T10</td>
<td>21/38</td>
<td>6/25</td>
<td>2.30</td>
<td>(0.82–6.50)</td>
<td>16/44</td>
<td>11/21</td>
<td>0.69</td>
<td>(0.27–1.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T11-L2</td>
<td>12/48</td>
<td>3/28</td>
<td>2.33</td>
<td>(0.61–8.99)</td>
<td>12/48</td>
<td>4/28</td>
<td>1.75</td>
<td>(0.51–5.95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trophic changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5–T7</td>
<td>24/35</td>
<td>10/21</td>
<td>1.44</td>
<td>(0.58–3.59)</td>
<td>27/31</td>
<td>15/17</td>
<td>0.99</td>
<td>(0.42–2.34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T8–T10</td>
<td>25/34</td>
<td>11/21</td>
<td>1.40</td>
<td>(0.57–3.43)</td>
<td>24/35</td>
<td>9/23</td>
<td>1.75</td>
<td>(0.69–4.44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T11-L2</td>
<td>19/40</td>
<td>9/23</td>
<td>1.21</td>
<td>(0.47–3.12)</td>
<td>22/38</td>
<td>8/24</td>
<td>1.74</td>
<td>(0.67–4.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tissue changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5–T7</td>
<td>38/21</td>
<td>19/13</td>
<td>1.24</td>
<td>(0.51–3.00)</td>
<td>36/23</td>
<td>17/15</td>
<td>1.38</td>
<td>(0.58–3.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T8–T10</td>
<td>42/17</td>
<td>21/11</td>
<td>1.29</td>
<td>(0.51–3.25)</td>
<td>47/12</td>
<td>20/12</td>
<td>2.35</td>
<td>(0.90–6.11)†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T11-L2</td>
<td>45/14</td>
<td>19/12</td>
<td>2.03</td>
<td>(0.79–5.19)</td>
<td>47/12</td>
<td>15/17</td>
<td>4.44</td>
<td>(1.73–11.37)‡</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Trade Offs

• Significance, power, meaningful effect size and sample size are directly related
• There are trade offs between them
• Researchers plan their studies looking for the correct balance between significance, power, and sample size to observe a meaningful effect