### Evaluating the Evidence

**P. W. Stone**  
M6728  
Columbia University, School of Nursing

---

### Objectives for this Session

1. Identify the types of evidence available, and their strengths.  
2. Review the structure of a research report.  
3. Identify strategies for evaluating validity and applicability of the evidence.  
4. Critique article and evaluate validity of evidence.

---

### Definition of Evidence Based Medicine/Practice

Conscientious, explicit & judicious use of current best evidence in making decision about care of individual patients  

David Sackett
Evidenced-Based Healthcare

The conscientious, explicit, and judicious integration of
- the best available evidence from systematic research,
- with individual clinical expertise and
- patient preference
at the bedside or in the clinic
to make decisions about clinical practice.

5 Step Evidenced-Based Process

1. Identify a focused clinical question (PICO).
2. Efficiently track down the best evidence to answer it.
3. Critically appraise that evidence.
4. Apply valid, useful evidence in your practice.
5. Evaluate your performance.

Overview of Finding, Obtaining and Evaluating the Evidence

1. Clinical Problem
2. Define an important, searchable question
3. Select most likely source
4. Design search strategy
5. Summarize the evidence yield

If adequate,
6. Apply the evidence
   Repeat, if necessary
Sackett et al. 1997
**Online Resources/Tutorials**

On-line at:
- [http://www.hsl.unc.edu/EBM/literat.htm](http://www.hsl.unc.edu/EBM/literat.htm)

Other EBM resources
- [cpmnet.columbia.edu/library/resources.html](http://cpmnet.columbia.edu/library/resources.html)
  - select EBM

**Finding the evidence**

*Select the most likely source*

- **Electronic databases**
  - Medline
  - CINAHL
  - Psychlit
- **Database software**
  - OVID
- **Web sites**
  - PubMed
  - Medline plus (for consumers)
- **Systematic Reviews of RCT**
  - EBM clinical evidence
  - Cochrane Collaboration
- **Electronic Journals**
  - Online Journal of Knowledge Synthesis for Nursing
- **Hand searching**

**How do I access databases?**

- [http://www.columbia.edu/cu/libraries](http://www.columbia.edu/cu/libraries)
- [http://www.infotrieve.com/freemedline](http://www.infotrieve.com/freemedline)
Why review the literature

- Get new ideas/approaches for practice
- Identify relevant studies/investigators
- Suggest methods
- Identify other sources of information
- Place your own study and practice in perspective
- Evaluate various studies by comparison

“Failure to search the appropriate scientific literature is an obvious breach of the broader duty to perform at the level of knowledge and practice in a ....clinical specialty”
Skolnick, Medicine and Law, 1985

Narrowing the field....

- specific population of interest
- place of interest (e.g. longterm care)
- discrete time period
- selected disciplines or journals
- selected experts in the field
- Use of EBP filters

PICO
**Strength of Evidence**

I Strong evidence from at least one systematic review of multiple well-designed RCTs

II Strong evidence from at least one properly designed RCT of appropriate size

III Evidence from well-designed trials without randomization, single group pre-post, cohort, time series or matched case-control studies

IV Evidence from well-designed non-experimental studies from more than one center or research group

V Opinions of respected authorities, based on clinical evidence, descriptive studies or reports of expert committees.

Source: Muir Gray, 1997

---

**Level 1**

- Systematic Reviews
  - Validity
    - How were the studies selected?
    - Are the studies included valid?
    - Are the results consistent across studies?
  - Applicability
    - How are the study participants similar or different from your patients?

- Meta-analysis
  - Combines statistical results of studies.
  - Critique is the same as systematic review.
## Level II

- **Validity of The RCT (internal validity)**
  - Sample size
  - If difference was found, could it be associated with another reason
    - Randomized subjects
    - All subjects accounted for
    - Blinded to treatment groups
  - Clinical versus statistical meaningfulness
- **Applicability of the RCT (generalization-external validity)**
  - How study participants compare to your patients
  - Your ability to give same treatment

## Level III

- **Non-randomized Experimental Designs**
  - Type of designs (quasi-experimental)
    - Matched case-control, single group pre-post, time series
  - Validity:
    - Was the comparison valid?
    - If a difference was found, could anything else of caused it?
    - Change overtime, groups not really comparable, researchers wanted to find the difference
- **Applicability**

## Level IV

- **Non-experimental designs (correlation studies) from more than one research center**
  - Validity
    - Do the different researchers find the same thing?
  - Applicability
Level V

- Expert Opinion
- Position Statements
- Be a careful and critical thinker as well as up-to-date
  - AHRQ patient safety review

Other rating systems

- Many groups have different (but similar) rating systems for ranking the level of evidence.
  - CDC
  - Cochrane

Nursing Journals

<table>
<thead>
<tr>
<th>Published study</th>
<th>1975</th>
<th>1985</th>
<th>1995</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only nurse authors:</td>
<td>66%</td>
<td>79%</td>
<td>75%</td>
</tr>
<tr>
<td>Doctorate:</td>
<td>16%</td>
<td>25%</td>
<td>80%</td>
</tr>
<tr>
<td>Mean # refs:</td>
<td>7.3</td>
<td>11.2</td>
<td>28</td>
</tr>
</tbody>
</table>
HINTS

- Start from specific and move to general
- Look for review articles
- Use reference lists in good papers
  - Seminal literature

General Format of Research Article

- Abstract
- Introduction
- Methods
- Results
- Discussion and Conclusions
- References

ABSTRACT

- Brief summary of why, who, how, what
- Usually about 150-250 words
## INTRODUCTION

- Statement of purpose and background
- Answers “so what?”
- Usually a few paragraphs

## METHODS

- Study design
- Setting
- Sample, sample size
- Procedures/definitions
- Instruments
- Analysis

## RESULTS

- Presents raw data first, then statistical analyses
- Specifically addresses each research question
- NO interpretation or discussion
- Usually 1-2 pages
Common Statistical Problems

- Inappropriate tests
- Assumption that significance = causality
- Insufficient sample size or time (lack of power)
- Assuming statistical significance = "good" study
- Confusing statistical and clinical significance

---

Power and Types of error

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Reality-No difference</th>
<th>Reality-Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision</td>
<td>Reality-No difference</td>
<td>Reality-Difference</td>
</tr>
<tr>
<td>Accept $H_0$</td>
<td>OK</td>
<td>Type II</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>Type I</td>
<td>OK (power)</td>
</tr>
</tbody>
</table>

**Type I:** Reality, there is no difference but they report one

- Risk of making a Type I error
- $\alpha$ (alpha)
- can be decreased by decreasing the level of significance (which decreases the power)
- Tradition-statistical significance is $p \leq 0.05$
**Type II:** Reality there is a difference, but they don’t find it!

- Need to **INCREASE POWER**
  - Make significance level less extreme (change from 0.01 to 0.05, or 0.10)
    - this is done in some pilot studies
  - Increase the sample size
  - Increase sensitivity of test to detect differences (decrease error in measurement)

**The trade off**

- Decreasing the probability of making a Type I error increases the probability of making a Type II error.

- What to do depends on the type of study

**P values**

- $p = 0.001$ means:
  - there is 1 in 1000 you are wrong and the difference actually occurred by chance.

The **probability** of committing a Type I error.

Type I- there is no difference but they report one
## DISCUSSION

- Author’s interpretation of results
- Comparison with other studies
- Possible explanations of findings
- Strengths and limitations
- Implications for practice and research
- Recommendations and conclusions
- Usually 2-4 pages

## Pitfalls in Conclusions

- Statistical vs. clinical significance
- Correlation vs. causality
- Biased language
- Biased citation of literature
- Lack of comparability with other studies
- Reader bias

## Watch for Potential Errors: Data Characteristics

- Inadequate sampling
- Inaccurate measurement
- Unrepresentative data
- Careless observation
- Intentionally distorted data
- Unclear definitions
### REFERENCES
- Style varies, depending upon journal

### General Criteria for Critique
- Clarity and relevance of purpose
- Researchability of the problem
- Adequacy of literature review
- Match of purpose, design, methods
- Suitability of sampling and sample
- Correctness of analytic strategy
- Clarity of findings

### Applying the evidence to practice
- Determining if the study is valid
  - Do you believe the results?
- Determining if valid results are applicable to your setting
  - Could you apply the evidence and expect the same results?
**Group Assignment**


- Complete the sample EBP worksheet

**Next Week Clinical Application 1 Due**

- Write an answerable question.

- Conduct literature search.
  - Write 1 paragraph on how literature was conducted.

- Select 2 of the best papers.
  - Attach copy of abstracts.

- In 1-2 paragraphs, synthesize the research
  - Includes assessing the reports strength of evidence (validity)
  - Includes assessing the reports applicability to your question.