Systematic Review & Meta-Analysis

Mary E. Lough PhD, RN, CCRN, CCNS, FCCM
Research Scientist, Clinical Nurse Specialist
Stanford Health Care
Clinical Assistant Professor
School of Medicine, Stanford University
Clinical Professor
Dept. of Physiological Nursing, UCSF

August 25th 2017
Conflict of Interest

- No conflicts of Interest
Objectives

1. Describe a reproducible method for your literature review
2. Discuss reasons to include systematic reviews in your literature review
3. State the purpose of a systematic review and meta-analysis
4. Interpret a Forest Plot (group work in the lecture)
5. Compare statistical significance and clinical usefulness in a systematic review and meta-analysis
Develop a reproducible method for your literature review

Clinical Question

Literature Review: What has been published?

Create an Evidence Table: Assess Quality

Literature Appraisal & Interpretation
What is your clinical question?

- PubMed
- CINAHL
- Google Scholar
- Google
- Cochrane review
- Joanna Briggs Institute (JBI)
- Check reference lists / recommendations

Search Publication Databases

Insert key words relative to your clinical question

Determine What has been published?
Identify the types of literature – Not all studies are equal!

- Meta Analysis & Systematic Review of RCTs (highest quality)
- RCT (single and multi center)
- Quasi experimental studies (has an intervention but not RCT)
  - Case control studies
  - Intervention with historical controls (pre-post design)
- Observational studies
- Case studies and case series
- Expert opinion (lowest quality in research terms)

See next slide for definitions of Quality
What is meant by “Quality” & “Rigor” in a Research Design

Questions to ask:

- How much control of the research protocol was present?
- In a randomized controlled trial (RCT) this describes **internal validity**
- How generalizable are the results to other settings? Known as **external validity**

https://www.cebma.org/faq/what-are-the-levels-of-evidence/

http://www.indiana.edu/~p1013447/dictionary/ext_val.htm
Randomized Controlled Trial (RCT):

- Patients meeting eligibility criteria who agree to be in a study
- Randomization procedures ensure patients are allocated by chance not bias
- Groups are equal at baseline

Group A

Group B

Confidential – For Discussion Purposes Only
Clinical Question

No random Assignment

There will be an intervention

Exposure to the intervention is not by chance

Groups may not be equal at baseline

Cannot assign 'causality'

Quasi-Experimental Trial: Pre-test with Post-test design

Pre-test

Post-test

1. Change over time
2. Learns answers
Quasi-Experimental Study (one type)

Historical Control: Patients who were in hospital in 2015 (Jan – Dec)

Quasi-Exp. research study:
1. No randomization
2. Groups may not be equal
3. Hospital practice may be different in 2015 than 2016

Intervention Introduced – January 2016

Current Group: Patients who were in hospital in 2016 (Jan – Dec)

The decision to do the study is often taken when the new intervention is introduced.

Retrospective chart review

Group A

Current patients

Group B
What do “Power” and “Effect Size” mean – and why do you care?

- **Statistical Power**:  
  - Statistical power is a measure of the likelihood that a researcher will find statistical significance in a sample if the effect exists in the full population.
  - Factors that impact statistical power
    - sample size
    - effect size
    - small 0.2; medium 0.5; large 0.8
    - significance level (statistical & clinical)
    - Statistic (effect measure)

**Literature Appraisal & Interpretation:**

Was there sufficient ‘power’ to answer the clinical question?

Many RCT sample sizes are too small to answer the question clinically

www.biochemia-medica.com/content/power-analysis-research
Evaluate the Evidence

- Use a **standardized method** to evaluate (Grade) the evidence.

- You have many choices ........
  - Cochrane - super-detailed with many sub-categories
  - John Hopkins – straightforward (*next 2 slides*)
  - Joanna Briggs Institute (JBI) (Australia)
  - Critical Appraisal Skills Program (CASP) checklists (UK)
  - Oxford EBM (UK)
# Literature Appraisal Tools – Johns Hopkins

Evidence Level and Quality example

## Johns Hopkins Nursing Evidence-Based Practice

### Appendix C: Evidence Level and Quality Guide

<table>
<thead>
<tr>
<th>Evidence Levels</th>
<th>Quality Guides</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level I</strong></td>
<td><strong>A High quality:</strong> Consistent, generalizable results; sufficient sample size for the study design; adequate control; definitive conclusions; consistent recommendations based on comprehensive literature review that includes thorough reference to scientific evidence</td>
</tr>
<tr>
<td>Experimental study, randomized controlled trial (RCT)</td>
<td><strong>B Good quality:</strong> Reasonably consistent results; sufficient sample size for the study design; some control, fairly definitive conclusions; reasonably consistent recommendations based on fairly comprehensive literature review that includes some reference to scientific evidence</td>
</tr>
<tr>
<td>Systematic review of RCTs, with or without meta-analysis</td>
<td><strong>C Low quality or major flaws:</strong> Little evidence with inconsistent results; insufficient sample size for the study design; conclusions cannot be drawn</td>
</tr>
<tr>
<td><strong>Level II</strong></td>
<td></td>
</tr>
<tr>
<td>Quasi-experimental study</td>
<td></td>
</tr>
<tr>
<td>Systematic review of a combination of RCTs and quasi-experimental, or quasi-experimental studies only, with or without meta-analysis</td>
<td></td>
</tr>
<tr>
<td><strong>Level III</strong></td>
<td></td>
</tr>
<tr>
<td>Non-experimental study</td>
<td></td>
</tr>
<tr>
<td>Systematic review of a combination of RCTs, quasi-experimental and non-experimental studies, or non-experimental studies only, with or without meta-analysis</td>
<td></td>
</tr>
<tr>
<td>Qualitative study or systematic review with or without a meta-synthesis</td>
<td>3 Levels</td>
</tr>
</tbody>
</table>

Literature Appraisal Tools: Johns Hopkins Evidence Summary Table

### Johns Hopkins Nursing Evidence-Based Practice
#### Appendix G: Individual Evidence Summary Tool

**EBP Question:**

**Date:**

<table>
<thead>
<tr>
<th>Article #</th>
<th>Author &amp; Date</th>
<th>Evidence Type</th>
<th>Sample, Sample Size &amp; Setting</th>
<th>Study findings that help answer the EBP question</th>
<th>Limitations</th>
<th>Evidence Level &amp; Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Attach a reference list with full citations of articles reviewed for this EBP question.

© The Johns Hopkins Hospital/Johns Hopkins University. May not be used or reprinted without permission.

Systematic Reviews & Meta Analysis

What you need to know

- Literature Review & Select Studies
- Systematic Review of selected studies
- Meta-analysis Forest Plot
Systematic Reviews & Meta Analysis

- Groups patients from similar randomized controlled trials (RCT) together to give more “power” for the statistical analysis.
- Important to assess how ‘similar’ the individual groups are
- Larger number of subjects in any study means a greater likelihood of detecting a statistical difference, if one exists.
- Small studies are sometimes described as “underpowered” because they do not have a sufficient sample size to detect a meaningful statistical difference between groups as determined by the p value.
- A meta-analysis combines studies to increase sample size and power
- Meta analysis uses a statistical technique called a Forest Plot to show outcome differences (variables) between groups: next slide.
What is a Forest Plot: Key Elements

#1 First Author and Year
#2 First Author and Year
#3 First Author and Year
#4 First Author and Year

Square indicates “weight” of study typically sample size

Diamond is the summary statistic

Line of no effect

Label of effect

Label of effect

1 indicates binary variable
0 indicates continuous variable

Binary Variable Interpretation:
If binary, confidence interval (CI) includes 1? ie $[0.98, 1.12] = \text{not significant}$
Do the “whiskers” cross 1? The arms reflect the CI value

Continuous Variable Interpretation:
If continuous CI range includes 0? ie $[0.42, 1.68] = \text{not significant}$
Do the “whiskers” cross 0? The whiskers reflect the CI value
1. Study Selection process: is it clear and without bias?

- Studies identified through database search (until May 10, 2016) n= 1536
- Studies (n= 1) identified from searching reference lists of reviews (n= 14)
- Articles screened based on title and abstract (n=953)
- Included (n= 18)
- Full copies retrieved and assessed for eligibility: studies meeting inclusion criteria
- Included (n= 10)
- Number of studies included in the systematic review (n= 9)
- Number of studies included in the meta-analysis (n= 7)

**Are the inclusion and exclusion criteria clearly explained?**

**How many people reviewed the studies?**

**Were the reviewers blinded?**

---


**Dichotomous outcome (Risk Ratio):** Line of No Effect is at 1.0

**Diamond = overall effect estimate**

**Statistical Significance:** Do the “whiskers” cross 1?

**Confidence Interval (CI) range:** Same as ‘whiskers’ in plot. If value crosses 1.0 the result is not significant.

‘Whiskers’ If line crosses 1.0 the result is not statistically significant.

What is being compared?

Study authors and year.

Weight of the study – usually sample size.
Meta analysis and Forest plot: Example #2

- Study Selection

Meta analysis and Forest plot: Example #2

Forest plot #1
PICC vs CVC by patient type only (in-patient or out-patient)

Forest plot #2
PICC vs CVC by patient type (in-pt & out-pt) plus number of catheter days

Same article/topic: different data inclusion produces different results

Meta Analysis and Forest Plot Example #3


### Clinical Significance

3.3.1. **HbA1c**

Of the total trials, 36 trials with 6920 participants reported on HbA1c, and showed a small but significant greater reduction for nurse-led interventions compared to ‘usual care’: [mean difference is $-0.28\%$; 95% CI: $-0.33\%$ to $-0.23\%$; p-value < 0.0001, Fig. 2]. Heterogeneity was high indicating large

### Statistical Significance

3.3.1. **HbA1c**

Of the total trials, 36 trials with 6920 participants reported on HbA1c, and showed a small but significant greater reduction for nurse-led interventions compared to ‘usual care’: [mean difference is $-0.28\%$; 95% CI: $-0.33\%$ to $-0.13\%$; p-value < 0.0001, Fig. 2]. Heterogeneity was high indicating large

### Statistical Interpretation

If the CI value includes 0, result is not significant

Example: [-1.95, 1.35]

If whisker crosses 0 (line of no effect) not significant

---

**Heterogeneity statistic (I²)**

$I^2 = 0\%$ no variation

$I^2 < 25\%$ low heterogeneity (more similar)

$I^2 = 50\%$ moderate heterogeneity

$I^2 > 75\%$ high heterogeneity (less similar)
Please work with a partner to interpret the following meta analysis and forest plots
Questions: Change in SBP, DBP & Triglycerides

Is this a binary variable or a continuous variable?

Did the nurse-led interventions work? How do you know?

How heterogeneous ($I^2$) are the studies?

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Nurse led intervention</th>
<th>Usual care (control)</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
</tr>
<tr>
<td>Blackberry 2013</td>
<td>133</td>
<td>14</td>
<td>189</td>
</tr>
<tr>
<td>Crowley 2013</td>
<td>137.6</td>
<td>17.5</td>
<td>182</td>
</tr>
<tr>
<td>Francosi 2011</td>
<td>133</td>
<td>13.5</td>
<td>146</td>
</tr>
<tr>
<td>Frei 2014</td>
<td>136.4</td>
<td>17.5</td>
<td>145</td>
</tr>
<tr>
<td>Frosch 2011</td>
<td>129.1</td>
<td>19</td>
<td>100</td>
</tr>
<tr>
<td>Gabbay 2006</td>
<td>129</td>
<td>18</td>
<td>150</td>
</tr>
<tr>
<td>Gabbay 2013</td>
<td>131</td>
<td>16</td>
<td>188</td>
</tr>
<tr>
<td>Jansink 2013</td>
<td>141.5</td>
<td>17</td>
<td>120</td>
</tr>
<tr>
<td>Jayasunya 2015</td>
<td>124.1</td>
<td>16</td>
<td>28</td>
</tr>
<tr>
<td>Ochoa-Catalina 2016</td>
<td>128</td>
<td>19</td>
<td>241</td>
</tr>
<tr>
<td>Piette 2011</td>
<td>130.8</td>
<td>17.7</td>
<td>145</td>
</tr>
<tr>
<td>Scal 2009</td>
<td>120.1</td>
<td>16.7</td>
<td>52</td>
</tr>
<tr>
<td>Schilling 2009</td>
<td>136.9</td>
<td>20.4</td>
<td>107</td>
</tr>
</tbody>
</table>
| Sheh 2009          | 136.12|20.38|262|139.46|22.22|373|9.1%|3.36 [-6.44, 3.28]

Total (95% CI): 2055
Heterogeneity: $I^2 = 8.37; \chi^2 = 38.77, df = 13 (P = 0.0002); FI = 66$
Test for overall effect: $Z = 1.86 (P = 0.06)$

Fig. 3 – Mean reduction in systolic blood pressure (mmHg).

Meta Analysis Examples – Please work in groups

Questions: Change in SBP, DBP & Triglycerides

Is this a binary variable or a continuous variable?

Did the nurse-led interventions work? How do you know?

How heterogeneous ($I^2$) are the studies?


**Fig. 4** – Mean reduction in diastolic blood pressure (mmHg).
Meta Analysis Examples – Please work in groups

Questions: Change in SBP, DBP & Triglycerides

Is this a binary variable or a continuous variable?

Did the nurse-led interventions work? How do you know?

How heterogeneous ($I^2$) are the studies?

### Questions:

**Meatal cleaning and CAUTI**

- Is this a binary variable or a continuous variable?
- Did meatal cleaning reduce CAUTI? How do you know?
- How heterogeneous ($I^2$) are the studies?
- Other comments?
  - Year of Publication
  - Clinical Applicability

---

Meta Analysis Examples – Please work in groups

Results:

Subgroup analyses were performed. Heterogeneity was estimated using the $I^2$ statistic. **Findings:** In total, 2665 potential papers were identified; of these, 14 studies were eligible for inclusion. There was no difference in the incidence of CAUTIs when comparing antiseptic and non-antiseptic agents (pooled OR 0.90, 95% CI 0.73–1.10; $P=0.31$), or when comparing different agents: povidone-iodine vs routine care; povidone-iodine vs soap and water; chlorhexidine vs water; povidone-iodine vs saline; povidone-iodine vs water; and green soap and water vs routine care ($P>0.05$ for all). Comparison of an antibacterial agent with routine care indicated near significance ($P=0.06$). There was no evidence of heterogeneity ($I^2=0%$; $P>0.05$). Subgroup analyses showed no difference in the incidence of CAUTIs in terms of country, setting, risk of bias, sex and frequency of administration.

References for Meta Analysis & Forest Plots

- References for Systematic Reviews & Forest plots
  - CEBM Oxford UK has a great review sheet for systematic reviews
    http://www.cebm.net/critical-appraisal/

- Excellent YouTube video on Forest Plots
  - http://www.bing.com/videos/search?q=forest+plots+utube&view=detail&mid=82BA9DB52C6313B6C5C182BA9DB52C6313B6C5C1&FORM=VIRE

Standards for Reporting of Systematic Reviews and Meta-Analysis (PRISMA)
www.prisma.statement.org

GRADE: Systematic reviews and meta-analyses of high quality RCTs are considered Level 1 evidence.
Questions

Contact information:
mlough@stanfordhealthcare.org