Critical Appraisal of Evidence

A Focus on Intervention/Treatment Studies
Bernadette Mazurek Melnyk, PhD, CPNP/PMHNP, FAAN, FAAN
Associate Vice President for Health Promotion
University Chief Wellness Officer
Dean and Professor, College of Nursing
Professor of Pediatrics & Psychiatry, College of Medicine
Editor, Worldviews on Evidence-Based Nursing

Common Beliefs about and Barriers to Critical Appraisal
- It takes too much time
- It is too difficult
- It takes a researcher or expert in EBP to do it well
- You have to be a “whiz” at statistics

Critical Appraisal: The Bottom Line
Being able to determine if findings from a valid study can be generalized to your patient(s) and factoring those findings into decision-making with your patients

Three Key Questions in Critical Appraisal
- Are the findings valid (as close to the truth as possible)?
- Are the findings important (reliable)
  - What is the impact of the intervention (size of the effect)?
- Are the findings clinically relevant/applicable to my patient(s)?

Critical Appraisal of an Intervention or Treatment Study
The purpose of this type of study is to determine whether a causal relationship exists between an intervention or treatment and an outcome

Three Major Criteria for Establishing Causality
- The independent variable (the treatment or intervention) must precede the dependent variable or outcome in time
- There must be a relationship between the independent and dependent variable
- This relationship can not be explained by other extraneous factors
Explanations for Differences Between Study Groups in Intervention Studies

- Extraneous/confounding variables
- Differences between how the groups were treated during the clinical trial
- The treatment or intervention worked!

Critical Appraisal of a Therapy or Intervention Trial: The Basics

Are the results valid?

Five Major Questions

1. Were subjects randomly assigned to treatment groups and was the random assignment concealed from the individuals enrolling subjects?

2. Was the follow-up sufficiently long to study the effects of the treatment and were all patients accounted for at the end of the study?

Critical Appraisal of a Therapy or Intervention Trial

3. Were patients analyzed in the group to which they were assigned?

4. Was the control group appropriate?

5. Were the instruments used to measure the outcome variables valid and reliable?

Are the results valid?

Additional Questions

- Were patients and providers kept blind to treatment?
- Were the groups treated equally aside from the experimental treatment?
- Were the groups equal at the beginning of the study?

What are the results of the study and are they important?

What was the magnitude of the treatment effect?
- Relative Risk Reduction (RRR)
- Absolute Risk Reduction (ARR)
- Number Needed to Treat (NNT)
- Number Needed to Harm (NNH)
- Effect sizes

How precise was the treatment effect?
- Confidence intervals

Relative Risk Reduction

- Proportion of risk for bad outcomes in the *intervention* group compared to the *unexposed* control group
Outcome of Smoking Cessation at Six Months Following Acupuncture Therapy or Placebo (sham acupuncture)

<table>
<thead>
<tr>
<th>EXPOSURE</th>
<th>OUTCOME</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes (Quit)</td>
<td>Yes</td>
<td>5</td>
</tr>
<tr>
<td>No (Not quit)</td>
<td>No</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>28</td>
</tr>
</tbody>
</table>

Relative Risk Reduction (RRR)

\[
RRR = \left( \frac{|R_e - R_u|}{R_u} \right) \times 100
\]

Re (EER) = \( \frac{5}{5+23} = .179 \)
Ru (CER) = \( \frac{4}{4+23} = .148 \)
RRR = \( \left( \frac{|.179 - .148|}{.148} \right) \times 100 = 21\% \)

Absolute Risk Reduction (ARR)

The absolute difference between the unexposed and exposed groups’ risks (i.e., occurrence in the unexposed/control group subtracted from the occurrence in the exposed/intervention group)

\[
ARR = R_e - R_u \times 100
\]

Re (EER) = \( \frac{5}{5+23} = .179 \)
Ru (CER) = \( \frac{4}{4+23} = .148 \)
ARR = \( (|.179 - .148|) \times 100 = 3\% \)

Number Needed to Treat (NNT)

The number of patients that need to be treated to achieve one additional favorable outcome.

\[
NNT = \frac{1}{ARR}
\]

NNT = \( \frac{1}{.03} = 33 \)
33 people would need to treated with acupuncture to promote one person’s smoking cessation

Number Needed to Harm (NNH)

The number of patients that need to be treated to achieve one negative outcome.

\[
ARI (\text{absolute risk increase}) = (R_e - R_u) \times 100
\]

\[
NNH = \frac{1}{ARI}
\]

ARI = \( (|.179 - .148|) \times 100 = 35 \)
NNH = \( \frac{1}{.149 - .178} = 35 \)
35 people would need to be treated with acupuncture to promote continuation of smoking
Effect Size

An estimate of how large the treatment effect is, that is how well the intervention worked in the experimental group in comparison to the control group.

The larger the effect size, the stronger are the experimental intervention’s effects.

Effect Size

- Calculated as the mean of the experimental group minus the mean of the control group divided by the pooled or average standard deviation.
- If found to be so large that it is unlikely to have occurred by chance, then we say it is statistically significant.

Effect Sizes for Interventions

.2 = Small effect
.5 = Medium effect
.8 = Large effect

Statistical Significance versus Clinical Meaningfulness

Statistical significance: the results obtained in a study are unlikely to be caused by chance.
Statistical significance is largely dependent upon power and the number of subjects in a study. The larger the sample, the greater the power and probability of detecting significant differences between study groups, even when effect sizes are small.

Randomized Controlled Trial

In teens, what is the effect of imagery versus deep breathing on state anxiety level?

20 = Experimental Group
20 = Control Group

Findings of the independent t-test revealed p=.15

Randomized Controlled Trial

\[
\frac{50 - 40}{20} = .5 \text{ Medium Effect}
\]

This is clinically meaningful but not statistically significant.
Important Take Home Message
When evaluating intervention trials, the focus should be on the magnitude of effects instead of statistically significant differences

P value
- The probability of an event given the assumption that there is no true difference between intervention and control
- Unlike CIs, p values do not tell us anything about the precision of the measures or the size of the effect

Confidence Intervals (CI)
- The range in which the real answer lies with a given degree of certainty (usually 95%)
- Describe an interval around the estimated effect point (you can get results within this interval, if you replicate their intervention)
- A 95% CI for the mean is sometimes loosely described as having a 95% probability of containing the true mean

Confidence Intervals (CI)
- 95% CIs = range in which can expect the “true” result to lie 95% of the time
- Meaning that in 1 in 20 studies, on average, the real value will lie outside the CI limits

Confidence Intervals (CI)
- The wider the CI, the less confident we are about the true result
- CI width is greatly dependent on sample size
- If CI for RR or OR contains 1, then the results are not statistically significant
- If CI for other statistics contain 0, then results are NS

How Precise is the Treatment Effect?
Confidence Interval: The precision of the effect
A NNT of 5 with a 95% confidence interval of 3-8 would mean that we have 95% confidence that the true NNT value is between 3 and 8
Will the Results From this Study Help Me in Caring for My Patients?

- Are the results applicable to my patient(s)?
- Were all clinically important outcomes considered?
- What are the risks and benefits of the treatment?
- Is the treatment feasible in my clinical setting?
- What are my patient’s values and expectations for both the outcome that is trying to be prevented and the treatment itself?

Exercise in Rapid Critical Appraisal of a RCT


Are the Study Findings Valid?

- Were the subjects randomly assigned to the experimental and control groups?
-Were the follow-up assessments conducted long enough to fully study the effects of the intervention?
- Did at least 80% of the subjects complete the study?

Are the Study Findings Valid?

- Was random assignment concealed from the individuals who were first enrolling subjects into the study?
- Were the subjects analyzed in the group to which they were randomly assigned?
- Was the control group appropriate?
What are the Results of the Study and are They Important?

- How large is the intervention or treatment effect (NNT, NNH, Effect Size, Level of Significance)?
- How precise is the intervention or treatment (Confidence interval)?

Will the Results Help Me in Caring for my Patients?

- Are the results applicable to my patients?
- Were all clinically important outcomes measured?
- What are the risks and benefits of the treatment?

Child Outcomes

<table>
<thead>
<tr>
<th></th>
<th>BASC (Externalizing Behaviors)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COPE</td>
</tr>
<tr>
<td></td>
<td>Effect Size*</td>
</tr>
<tr>
<td>n</td>
<td>M</td>
</tr>
<tr>
<td>Time 4</td>
<td>73 98.6</td>
</tr>
<tr>
<td>Time 7</td>
<td>49 99.6</td>
</tr>
<tr>
<td>Time 8</td>
<td>55 100.0</td>
</tr>
<tr>
<td>Time 9</td>
<td>43 97.3</td>
</tr>
</tbody>
</table>

*Small effect=.2; Medium=.5; Large=.8
**p<.05

Helpful References on Critical Appraisal


Contact Information

bernmelnyk@gmail.com

614-292-4844

Follow me on Twitter @bernmelnyk

Slide Presentation Copyright, 2015