Critically Reviewing of a Research Article

Learning Objectives:

1. Describe the approach to reviewing a paper
2. Identify the research hypothesis
3. Identify the quality of the research and the validity of the findings
4. Describe the factors which may raise concern about the truth of a research finding

Reviewing a Paper is about Asking Questions

Key Areas of Focus

- Critique selected population/sample selected
- Critique method of assessment
- Critique how the data was interpreted

What to Examine When Evaluating Data Collection

- Study Objectives
  - Is a hypothesis stated?
- Exposure and Outcome Variables
- Study Design
- Study Population
- Potential for
  - Selection Bias
  - Information Bias
  - Confounding

Study Context

- Several Issues to consider
  - What is the public health/clinical importance or significance of this study?
  - Does this study generate new hypotheses or confirm previous results with improved methods?
  - **Is the study hypothesis biologically plausible? (very important)**

Study Objectives

- What do the investigators want to achieve in this research?
  - What is the hypothesis tested in the study
    - Is the hypothesis specific or too general to refute?

Exposure and Outcome Variables

- Primary exposure
  - How was variable defined?

  Ex: Dental caries (DMFT)
- How was information on exposure collected?
- Records, Interviews, qestionnaires
  - Best method?
  - Sensitivity/specificity of this method?
  - Potential for misclassification?

- Primary outcome
  - Conceptual vs. operational outcome?
  - Oral Health Related quality of life vs Oral Health Impact Profile
- How was information on outcome collected?
  - Best method?
  - Sensitivity/specificity of this method?
  - Potential for misclassification?

Type of Study

- What study design was employed?
- Is it an appropriate design?
  - Exposure or outcome rare?
  - New hypothesis?
- What are the limitations and strengths of this design?

Type of Study

- What study design was employed?
- Is it an appropriate design?
  - Exposure or outcome rare? Neither
  - New hypothesis? No, but conflicting study results
- What are the limitations and strengths of this design?
  - Strengths: longitudinal assessment, incidence of dementia, uses previously collected data
  - Limitations: short period of assessment

The Study Population

- What was the source of study population?
- Hospitals, school, phone book
  - How does the study population compare to the general population?
  - Will the findings relates to my patients.
- How were subjects selected?
  - Could this method introduce selection bias?
  - Dentists/doctors selection of patients.
- What was the sample size?
  - Is the statistical power of the study identified?

The Study Population

- What was the source of study population?
- How were subjects selected? Not stated
  - Could this method introduce selection bias?
- What was the sample size?
  - Is the statistical power of the study identified (80%, 90%)? Yes
  - Out of the projected study sample, how many persons participated?
  - Ex: What was the response rate????

Potential for Confounding

- What factors were potentially confounding the study relationship?
- What methods did the authors use to minimize the influence of confounding when planning the study?
  - E.g. restriction, matching, randomization, etc.

What to Examine When Evaluating Data Analysis

- Confounding
- Measures of Association
- Measures of Statistical Stability

Data Analysis – Measures of Association

- What Measures of Association were reported in the study? Was the correct measure used?
- Cohort Study: Relative Risk (RR), Odds Ratio (OR), Hazard Ratio (HR), Incidence Rate Ratio (IRR.
- Case-Control: Odds Raio
- Cross-sectional Study: Prevalence Ratio.

Data Analysis – Statistical Stability

- Hypothesis Testing: Can use p-values or confidence Intervals (CI) to test the null hypothesis.

- P-value: The probability of observing the study results given that the null hypothesis is true. $P<0.05$ is a standard value that investigators use to reject the null hypothesis of no association and declare that there is a significant relationship between 2 variables.

Data Analysis – Statistical Stability

- 95% CI: This measure can be used for hypothesis testing and interval estimation. Can be defined as, if one will repeat the study 100 times the true association will lie inside the interval 95% of the time.

- We fail to reject the null hypothesis when a confidence interval contains the null value of 1 between its lower and upper limits for relative measures. EX: 95% CI: 0.8-4.5, It includes the value 1

Data Analysis – Statistical Stability

- Large confidence intervals indicate that the standard error is high. A high standard error is often related to a small sample size. Underpowered studies (less than 80%) normally have wider confidence intervals and thus difficulty in rejecting the null hypothesis.

- The problem, therein, lies that it is difficult to know if the non-association is real or false.

What to Examine When Interpreting the Results of the Study

- Major findings of the research
- Influence (on the results) of:
  - Bias and confounding
  - misclassification

Major Findings

- The first paragraph of the discussion section in a manuscript should summarize the main findings of the study.

Influence of Bias and Confounding

- Reader should be able to recognize information bias, selection bias, or confounding in the study and assess their magnitude and direction in the study (underestimate or overestimate the effect) .

- Bias or confounding that is large in magnitude indicates that the findings in this sample may not approximate what you would expect to see in the population.

Misclassification

- Misclassification of the exposure or the outcome (or both) can influence study results

Non-Differential: Misclassification is similar in the exposure or outcome groups. This would bias the results to the null (towards no difference) making it unlikely for investigators to reject the null hypothesis.

- Differential: Misclassification occurs at a different rate in exposure or outcome groups.

- Example of differential misclassification, a larger number of patients are classified as high anxiety instead of medium anxiety than individuals classified as medium anxiety instead of high anxiety. This type of misclassification can bias results away (find a large difference) or towards the null hypothesis (no difference).

Formulating an Overall Impression of the Manuscript

- What are the strengths and limitations of the paper?
- Can the results be generalized to the whole population?

Strengths and Limitations

- Examine the overall issues related to data collection, data analysis, and data interpretation.
- What conclusions do you draw from the results based upon your interpretation of the strengths and limitations of the study?
- Do the strengths outweigh the limitations?
- They are often mentioned in the discussion section of a manuscript.

Generalizability

- Goal is to have a study where the results can be used to infer what is going on in a certain population
- Major problems with the internal validity of the study make it difficult to for the results to be generalized to any population.
- Example, the study population excluded a certain groups, women, older people or low income individuals. Then the results would not be generalizable to these groups.

Conclusions

- The conclusions are a brief summary of the findings.
- Authors tend to include recommendations for future studies or policy.
- .

The Big Picture of Research Findings

- Publication bias
- John P.A. Ioannidis
  - Why Most Published Research Findings are False. PLoS Medicine 2(8):e124, 2005.

Statistical Significance

A finding is statistically significant when it allows the researcher to reject the null hypothesis at a pre-specified level of confidence or probability. This level is called the alpha level (α) and for educational purposes is usually set at .05.

Statistical Significance

Findings are said to be statistically significant when the *p* value associated with the test statistic is smaller than the predetermined alpha level (usually .05).

A *p* value of less than .05 would result in rejecting the null hypothesis.

Practical Significance

- While a finding may be statistically significant it may or may not be practically significant.
- A researcher must contemplate the usefulness of the finding to determine its practical significance.

Practical Significance

- When determining practical significance the researcher must consider the following:
  - The quality of the research questions
  - The relative size of the effect
  - The size of the sample
  - The importance of the finding
  - Confidence intervals
  - The link to previous research
  - The strength of correlation

Consider this example

In a study of 1000 children who received 1 mg of fluoride ion daily had a DMFS score of 1.5 while 1000 children who did not receive the fluoride had a DMFS score of 1.85. This difference was statistically significant.

What do you think of these results????

Would you recommend implementing of a fluoride tablet program

NO

While the 0.35 difference in DMFS may be statistically significant, it is not practically or clinically significant.

Statistical Significance and Clinical Importance

- Were both statistical and clinical significance considered?
- Significant results in a large study may sometimes not be clinically important.
- A greater problem arises from misinterpretation of non-significant findings.

Assessing the role of chance

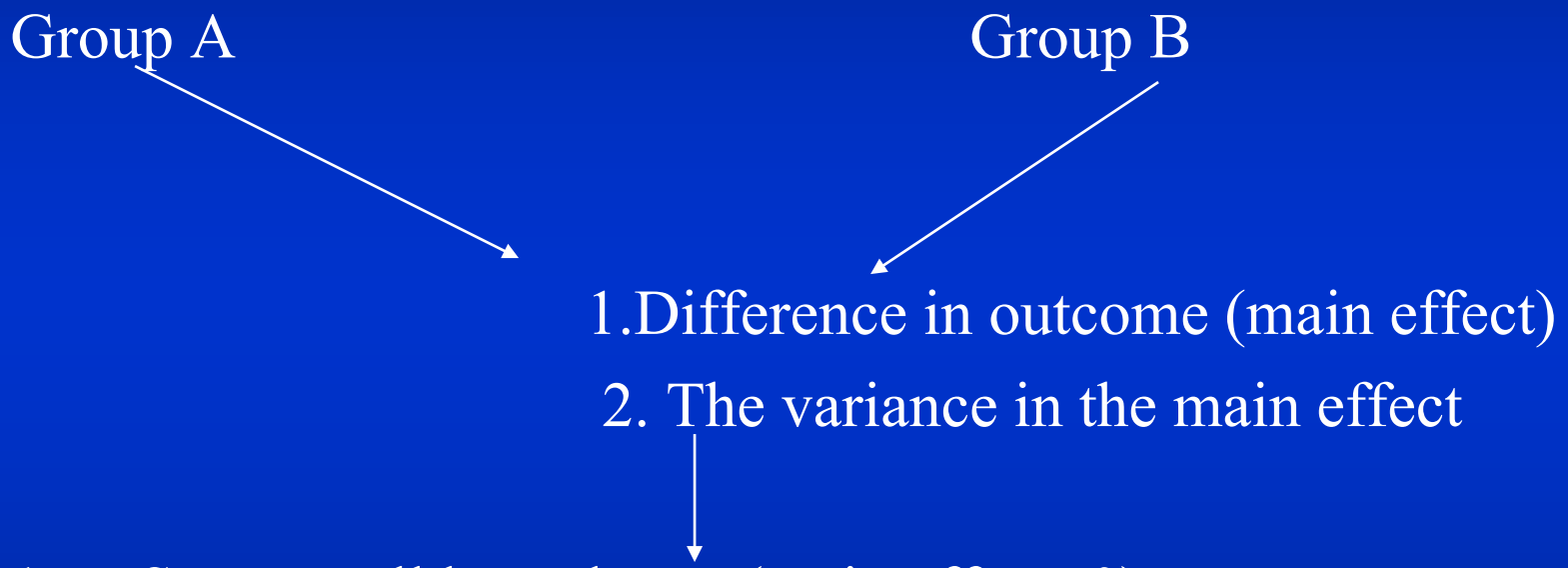
- Consider a study of a new weight loss program
- Group A receives te interventiona nd loses an average of 10 pounds
- Group B serves as a control group and loses on average 3 pounds

- There are two methods to assess the role of chance:
- Hypotheses Testing
- Confidence intervals

Hypotheses testing

- Group A
- Receives intervention and loses on average 10 pounds

- Group B
- Control group, loses on average 3 pounds



1. State a null hypotheses (main effect=0)
2. Calculate the P-value and 95% CI around the main effect

- The test address the question :
- If the true state of affairs is no difference (I.e the intervention does not lead to weight loss) THEN
- What is the probability of observing this difference (7 pounds) or more (8 or 9 pounds) by chance.

- Therefore: If the P value less than 5% researchers typically assert that the findings are “statistically significant”. Common language: UNLIKELY DUE TO CHANCE”.
- IF the P-value more than 5% then the findings are “statistically not significant”. Common language Likely to be due to chance
-

Factors that influence the P-value

- The magnitude of the difference: A 7 pounds difference will have a more significant results than a 1 pound difference.
- The number of observations:A 7-pounds difference observed in a study of 500 patients in each group will have a lower P-value than a 7 pounds difference in a study with 25 patients in each group.

- The spread in the data: If everybody in group A loses about 10 pounds and evryone in group B loses about 3 pounds, results will be more significant than if there is a wide variation in individual weight changes.

95% Confidence Intervals

- Using the same example, the average weight loss was 10 pounds in the inetrvention group and 3 pounds in the control group.
- Resulting in a mean difference of 7 punds and a 95% CI of 2-12.

- In other words, if the study was repeated 100 times, 95 times out of the 100 the CI will contain the truth and in 5 it will not.

Statistical significance does not translate into clinical importance

- Large studies can be significant without being clinically important.

N in each group	Weight loss Group A	Weight loss Group B	Main effect	P-value	Conclusion
10	20 lb	3 lb	17 lb	0.07	Not significant
1000	5lb	3 lb	2 lb	0.03	Significant, but clinically unimportant

Clinical Significance

Clinical significance refers to the practical or applied value or importance of the effect of the intervention--that is, whether the intervention makes a real (e.g., genuine,, practical, noticeable) difference in everyday life to the clients/patients or to others with whom the client/patient interacts.

Publication bias

- Definition:

"Publication bias refers to the greater likelihood that studies with positive results will be published"

JAMA 2002;287:2825-2828

Publication bias may ….

- Distort the scientific record
- Hide the “truth” of association/no association
- Influence doctors’ decision making
- Mislead policy makers
- Etc.

- The smaller the studies conducted, the less likely the research findings are to be true
- The smaller the effect size, the less likely the research findings are to be true
- The greater the financial interest and prejudice, the less likely the research findings are to be true
- The hotter a topic interest, the less likely the research findings are to be true

References:

Thompson, B. (Winter, 2002). “Statistical,” “Practical,” and “Clinical”: How Many Kinds of Significance Do Counselors Need to Consider? *Journal of Counseling and Development.* 80 p. 64 - 71.

http://ebp.lib.uic.edu/dentistry

Medical statistics : a guide to data analysis and critical appraisal (1st ed.) (2005)
Peat, Jennifer K.; Barton, Belinda.
ISBN: 9780727918123