AN INTRODUCTION TO CRITICAL APPRAISAL
PART 1

Dr. Ghadeer Aljuraiban

WHAT IS CRITICAL APPRAISAL?

- Research involves gathering of data, collection of data and analysis of the data to produce meaningful information.
- However, many of the research are not in good quality and many studies are biased and their results are untrue.
- This can lead us to draw false conclusions.

WHAT IS CRITICAL APPRAISAL?

- It is the process of carefully and systematically analyzing the research paper to judge its trustworthiness, and its value and relevance in a particular context.
- Balanced assessment of the benefits/strengths and flaws/weaknesses of a study
- Assessment of research process and results
- Consideration of quantitative and qualitative aspects

Hill and Spittlehouse, 2001

CRITICAL APPRAISAL IS NOT

Negative dismissal of any piece of research

Assessment on results alone

Based entirely on statistical analysis

Undertaken by experts only

WHY CRITICALLY APPRAISE?

Critical Appraisal aims to help people develop the necessary **skills** to make sense of scientific evidence based on validity, results and relevance:

To find out the **validity** of the study

- are the methods robust?

To find out the **reliability** of the study

- what are the results and are they credible?

To find out the **applicability** of the study

- is it important enough to change my practice?

HOW DO I CRITICALLY APPRAISE THE RESEARCH?

WHAT DO I NEED TO KNOW?

Awareness of study designs

Levels of evidence

Statistics!!

Checklists

Resources

READ THE ABSTRACT

- Are your issues discussed there?
- What are the main findings of the research?
- Do you want to know more after reading the abstract?
- Does it address a related question?
- Are there reasons to doubt the findings without reading the whole article?

METHODOLOGY SECTION

- The Methodology will give you a step-by-step description of exactly how the study was carried out.
- Where the study was done?
- From whom the data was collected ?
- Is it primary or secondary data
- And how the data was collected?

METHODOLOGY SECTION

- How good is the data?
- Does the study adequately control for differences between the groups being compared?
- Are the statistical methods appropriate?
- Is the sample large enough to produce significant results?

How good are the measures?

- Do the measures accurately reflect what the researcher was trying to measure (validity)?
- How clear and appropriate are these measures? (Too broad? Too narrow? Ambiguous?)
- Are the measures well established in either prior research or through pilot testing by the researcher, or are they ad hoc?

WHAT ARE THE AUTHOR'S CONCLUSIONS?

Compare the abstract to the Discussion

- The discussion section is more detailed and precise than the abstract, and will explain the limitations of the research and possible implications which are not mentioned in the abstract.

- **Compare the raw data given in the tables with the results analyzed in the discussion and conclusions**

  - Are the results reported in the conclusions consistent with what is reported in the tables?

  - Is the interpretation consistent with what the actual findings were?

- **How well are the results related to other research on the same topic?**

  - In the discussion or conclusions section, is there a review of how these results compare or contrast with prior research?

  - If this report found something different from previous research, then it's important to question on appraising the reliability of the findings.

AN INTRODUCTION TO CRITICAL APPRAISAL
PART 2

Dr. Ghadeer Aljuraiban

CRITICAL APPRAISAL OF AN ARTICLE ON HARM / RISK (COHORT STUDY)

- **Are the results of the study valid?**
- Primary Guides
  - Was the exposure status clearly defined and measured?
  - How was the outcome of interest measured?
  - What was the follow up time and was it adequate to measure the outcome of interest?
  - Was the outcome measured in the same way for both exposed & not exposed?
- Secondary Guides
  - How much was the attrition? (Loss to follow up)
  - Are there confounders that the investigator did not address?

- **What were the results?**
  - How strong is the association between exposure and outcome? (Look for Relative Risk, Hazard ratio)
  - How precise is the estimates of the risk ? (Look for 95% CI and p value for statistical test of significance)

- **Will the results help me?**
  - Are the results applicable to my population?
  - What is the magnitude of the risk ?

SELECT AND ASSESS STUDIES

Eligibility criteria for study selection can be applied

More than one reviewer can help reduce bias

Checklists/scoring systems

WHAT DO THE FINDINGS MEAN?

Effect measures – odds ratios, relative risk, mean difference

P-values

Confidence intervals

USING STATISTICS

Assess the weight of the evidence that a treatment works (or doesn't)

Give an estimate (and likely range) of the treatment effect

Test to see how likely it is that this effect would have been seen by chance

ODDS RATIO (OR)

Expresses the odds of having an event compared with not having an event in two different groups

$$OR = \frac{\text{odds in the treated group}}{\text{odds in the control group}}$$

- OR=1 treatment has identical effect to control
- OR<1 event is less likely to happen than not (i.e. the treatment reduces the chance of having the event)
- OR>1 event is more likely to happen than not (increases the chances of having the event)

Clinical trials typically look for treatments which reduce event rates, and which have odds ratios of less than one

IMPORTANCE OF DEFINING THE OUTCOME

Type of outcome		
Value of OR/RR	**Adverse outcome (e.g. death)**	**Beneficial outcome (e.g. stopped smoking)**
<1	New intervention better	New intervention worse
1	New intervention no better/no worse	New intervention no better/no worse
>1	New intervention worse	New intervention better

P-VALUES

The probability of finding the observed, or more extreme, results when the **null hypothesis (H_0)** of a study question is true

P-value results range from 0 to 1

The closer the p-value is to zero, the less chance there is that the effects of two interventions are the same

STATISTICAL SIGNIFICANCE

In general, p-values of either 0.05 or 0.01 are used as a cut-off value, although this value is arbitrary

Results larger than the cut-off are considered likely to attribute the event to chance, while results smaller than the cut-off value are likely to have occurred because of a real explanation (i.e. the result is less likely due to chance)

P-value of <0.05 indicates the result is unlikely to be due to chance,

P-value of >0.05 indicates the result might have occurred by chance.

BE CAREFUL…

A p-value in the non-significant range tells you that either there is no difference between the groups *or* there were too few subjects to demonstrate such a difference (ideally need to report confidence intervals)

There is not much difference between p=0.049 and p=0.051

P-values do not indicate the magnitude of the observed difference between treatments that is needed to determine the clinical significance

INTERPRETATION OF CONFIDENCE INTERVALS

Confidence interval is the range within which we have a measure of certainty that the true population value lies

OR

The confidence interval around a result obtained from a study sample (point estimate) indicates the range of values within which there is a specific certainty (usually 95%) that the true population value for that result lies.

(MeReC Briefing 2005)

WHAT CAN A CI TELL US?

Tells us whether the result is significant or not

The width of the interval indicates precision. Wider intervals suggest less precision

Shows whether the strength of the evidence is strong or weak.

The general confidence level is 95%. Therefore, the 95% CI is the range within which we are 95% certain that the true population value lies

CONFIDENCE INTERVALS REPORTED ON RATIOS (ODDS RATIO, ETC)

The ‘line of no effect’ centres around 1

If a CI for an RR or OR includes 1 (the line of no effect) then we are unable to demonstrate statistically significant difference between the two groups

ADVANTAGES OF A SYSTEMATIC REVIEW/META-ANALYSIS

Limits bias in identifying and excluding studies

Objective

Good quality evidence, more reliable and accurate conclusions

Added power by synthesising individual study results

Control over the volume of literature

DRAWBACKS TO SYSTEMATIC REVIEWS/META-ANALYSES

Can be done badly

- 2 systematic reviews on same topic can have different conclusions

Inappropriate aggregation of studies

A meta-analysis is only as good as the papers included

Tend to look at ‘broad questions’ that may not be immediately applicable to individual patients

CONCLUSION

Critical appraisal of systematic reviews and other research is well within your capabilities

Use a recognised checklist (eg CASP)

Update your literature searching skills regularly

THANK YOU FOR LISTENING