Introduction to Critical Appraisal of a Quantitative Paper

Student & Library Services
Learning and Research Support
Aim of workshop

• To consider one model for appraising quantitative research (based on reliability, internal and external validity)
• To practise using this model on a sample article
Validity in Quantitative Research

1. Internal
   • Does the research measure what it is supposed to be measuring?

2. External
   • Can the results be applied to the wider population?
Reliability

• How consistent are the measurements?

• Reliability concerned with random, one-off, errors whereas validity concerned with systematic or constant error –
  • e.g. improperly calibrated scales might be reliable, but would be invalid
Writing up the appraisal

• Go through each element one at a time e.g. randomisation or sampling approach and directly compare all your articles under this.

• Then move onto the next element.

• Are some better than others in terms of internal and external validity and reliability?

• After all the elements, can you say if any of your articles are better overall?
Strategies for internal validity

• Randomisation of sample into different groups (in RCT’s)
• Blinding – patient/clinician/researcher not aware which group patient is in - single/double/treble
• Similarity of groups at baseline
  • look for tables that compare the demographic details of groups
• Equal treatment of the groups
• Low attrition
• Power analysis
Strategies for external validity

• Appropriate type of sampling strategy used
• Sufficient size of sample
• Everyone who started the study was accounted for at the end of the study
• SURVEY ONLY... Response rate
<table>
<thead>
<tr>
<th>Type of sample</th>
<th>How it works</th>
<th>When to use and potential flaws</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convenience / Opportunity</td>
<td>Participants are selected from a group who happen to be available (e.g. patients attending GP surgery on Tues am).</td>
<td>Should be avoided if possible – most likely to be biased</td>
</tr>
<tr>
<td>Random</td>
<td>A target group is identified, and a random selection of people from that group is invited to participate - eg diabetic people</td>
<td>Use in studies to reflect average viewpoint of a population – least biased.</td>
</tr>
<tr>
<td>Stratified random</td>
<td>As random sample but the target group is first stratified according to a particular characteristic(s) – eg diabetic people on tablets.</td>
<td>Use when the target group is likely to have systematic differences by subgroup – subgroups need to be varied</td>
</tr>
<tr>
<td>Type of sample</td>
<td>How it works</td>
<td>When to use and potential flaws</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td><strong>Quota</strong></td>
<td>Participants who match the wider population are identified (e.g. into groups such as social class and gender age). Researchers are given a set number within each group to interview (e.g. so many young middle-class women)</td>
<td>To reflect outcomes as closely representative of the wider population as possible. Frequently used in political opinion polls, etc. May be bias within each group.</td>
</tr>
<tr>
<td><strong>Snowball</strong></td>
<td>Participants are recruited, and asked to identify other similar people to take part in the research.</td>
<td>Helpful when working with hard-to-reach groups. May be biased.</td>
</tr>
</tbody>
</table>
Strategies for reliability

• What has been done to ensure that the research is consistent?
  • Standardised questionnaires
  • Pilot questionnaires
  • Checks
Reliability reported in studies

- Reliability coefficient reported as close to 1
- **Stability** – produce same results with repeated testing
  - Test-retest
    - Same test administered more than once
  - Parallel forms
    - Two comparable versions of test exist
- **Internal consistency/homogeneity** – all items measure same concept
  - Cronbach’s alpha above 0.7
- **Equivalence** - same results
  - If parallel procedures used
  - Two or more observers
Measuring tools

• Eg questionnaire, Weight/height, BMI
  • Recognised?
  • How tested?

• If well-recognised instrument, validity and reliability might be only mentioned briefly with reference to previous studies
A new questionnaire / tool

• If a new questionnaire has been developed
• Expect to see in the paper discussion around various forms of validity
  • **Face validity** – expert panel / literature / pilot test
  • **Construct validity** - correlates with other measures that are presumed to measure the same construct
  • **Criterion validity** - uses an already existing **standardised** measure, with which to compare the performance of the new measure 0.7 or above ...

• Reliability
  • **Cronbachs alpha** – 0.7 or above = internally reliable
  • **Test re test** - Correlation coefficient should be 0.7 or above= stability of tool
References
