Critical Appraisal
Notes

Aim: To decide how useful a paper is for clinical decision making, directing future research or policy. Overall in evaluating the initial quality of a paper five key questions need to be asked

: Is it of interest?
: Why was it done?
: How was it done?
: What were the results?
: What are the implications?

Why. Observed measures of effect in studies may be causal or due to reverse causality, bias, confounding, random error or chance.

Techniques focus upon: internal validity, reported results and generalisability.

Thus objectives are to

1) Identify potential sources of bias: selection, information, performance, attrition, ascertainment.
2) Analyse handling of potential confounders and also consider potential impact of others not identified
3) Critique analysis and results, especially raw data, number and appropriateness of analyses. Are statistically significant results clinically or socially significant?
4) Understand author's implications and those of your own.

Many criteria for appraisal apply to all research methods, but others are method specific. For this reason checklists have been suggested. However in order to do this the research method be identified..

Descriptive
Survey, ecological

Analytical
Case-control, Cohort

Intervention
(Semi) experimental, trials

The checklists here are adapted from Crombie (The Pocket Guide to critical appraisal. BMJ publishing 1996) and can lead to an informed decision regarding a paper. However, although they can lead to an identification of a studies flaws there are several caveats

Nearly all papers contain some flaws. Research is difficult to design and harder to do. Subjects are constantly getting the way of the perfect study. The important point is to ask whether the flaws matter and affect the conclusions of a study

Do not simply count the flaws. A small number of limited flaws may be less important than one big one

Consider alternative scenarios (sensitivity analyses) e.g. what would be the worst case if all non-responders to a survey questionnaire were of one type. Some clinical trials anticipate this with intention to treat analyses. This can help identify the likelihood of a flaws impact on conclusions.

Standard Appraisal Questions

1) Are the aims clearly stated?
   If clear and focused hypotheses are stated in advance it is less likely that there was data trawling and easier to follow
2) Was the sample size justified?
   Were there enough subjects to meet the aims e.g. find the size of effect sought. Small trials could fail to find clinically important effects.
3) Are the measurements valid and reliable?
   Many measurements are subjective leading to the possibility of bias. Instruments should ideally have been tested for validity and reliability in the study population. Even "biological" measures can be systematically flawed e.g. blood pressure measurement.
4) Are the statistical methods described?
   Again these should be specified before a study to prevent later multiple testing which can lead to spurious "significant" results. The statistics should start with simple analyses and only proceed to complex analyses if justified
5) Were there untoward events during the study?
   Attrition rates, missing data or data collected in different manners can lead to bias
6) Are the raw data adequately described?
   This is vital to allow readers to decide upon applicability in different settings
7) Do the numbers add up?
   Often participants are "lost" during a study but this should be explained. Small inconsistencies may not matter and reflect sloppiness but large should alert the careful reader
8) Are the statistical methods appropriate and properly executed?
   Are skewed data handled correctly. Are appropriate statistical tests of hypotheses (with p values) or estimates of precision (with confidence intervals) used? Increasingly journals prefer the latter.
9) What do the findings mean and were any overlooked?
   The size of observed effects, and the corresponding tests, should have been scrutinised. Bradford-Hill's "strength of evidence" for causality may be useful. This involves making inferences and is ultimately a matter a of judgement.
10) How are null findings interpreted?
    This could have been due to study size or design, possibly seen in wide confidence intervals. Lack of evidence of an association is not the same as no association
11) How do results compare with previous studies?
    This is difficult if not your area of expertise. Researchers often fit results in with their own views. It has been suggested that real advancement is made when people are able to look outside their paradigms.
12) Implications for you and others?
    Basically how big was the effect seen and what is it's importance, clinically, socially and/or politically. To what extent are the results generalisable?

Further method specific checklists: Is the design appropriate to the stated objectives?

1) Surveys

- Who was studied
- How was the sample obtained
- What was the response rate?

- Is there a suggestion of "study before design"?
- What about selection bias?
- Were the findings serendipitous?

- Can the results be generalised?

2) Cohort studies

- Who exactly has been studied
- Were controls used and the numbers of them necessary?
- How adequate was the follow-up?

- Was exposure adequately measured?
- Were all relevant outcomes assessed?
- Did the analysis allow for the passage of time?

- What else might influence outcome?

3) Case-control studies

- How were the cases obtained?
- Is the control groups appropriate?
- Was data collected the same way for cases and controls?

- Was there data dredging?
- What are the biases?
- Could there be confounding?

4) Intervention studies (trials)

- Were treatments randomly allocated
- What was allocation procedure
- Except for experimental treatments were groups treated equally?
- Are treatments adequately described to allow replication
- Were all patients accounted for?
- Was compliance assessed?
- Were outcomes assessed blind?
- Are outcomes clinically relevant

- Were treatment groups comparable at baseline?
- Were results analysed by Intention To Treat.?
- Were negative effects of treatment reported?