Peer review & Critical Appraisal

Dr. Gholamreza Khalili
Department of Epidemiology & Biostatistics
School of public health
Tehran University of Medical Sciences
The science of ‘trashing’ a paper

Unimportant
issue

Unoriginal

Hypothesis not tested

Different type of study
required

Compromised
original protocol

Sample size too small

Badly written

Conflict of interest

Unjustified
conclusion

Poor statistics
Peer review

- Articles submitted to peer-reviewed journals (manuscripts) are reviewed by experts who advise the editor on whether they should be published and what changes are necessary.
Peer Review - Functions

- To Protect
  i) The author from publishing &
  ii) The subscriber from reading

Material of insufficient quality
Editorial decision

An editorial committee may decide that a paper:

- Is acceptable for publication
- Is acceptable for publication following minor revisions
- Is acceptable for publication following major revision
- May be reconsidered for publication following major revisions
- May be considered for publication as a letter or a short report
- Is unacceptable for publication
Editorial decision

- Rejection rate: 15% (pay journals) to 60% (specialist journals) to 90% (NEJM, The Lancet)

- How long does it take? (Choice of journal)
  - BMJ: 70 days
  - JAMA: 117 days
  - Iranian journals?
Questions journals ask

- Is the research question important?
- Is it interesting to our readers?
- Is it valid? A scientifically sound study.
What editors look for

- Short, clear, precise title
- Good abstract
- Good design and methods
- Clear conclusions
- Brevity
- Follow instructions
What reviewers look for

- Good design and methods
- Simple tables and figures
- Logical organisation
- Brevity
- Balance
- Appropriate statistics
- Their papers
Problems with peer review

- Slow
- Expensive
- A lottery
- Ineffective
- Biased
- Easily abused
- Can’t detect fraud
Critical Appraisal Skills Programme (CASP)

Critical appraisal is the process of weighing up evidence to see how useful it is in decision making.
Critical appraisal helps the reader of research……

• Decide how trustworthy a piece of research is (validity)

• Determine what it is telling us (results)

• Weigh up how useful the research will be (relevance)
Critical Appraisal:
Three preliminary questions

- Why was the study done and what hypothesis was being tested?
- What type of study was done?
- Was the study design appropriate?
Why was the study done?

i.e. what was the key research question/ what hypotheses were the author testing?

“null hypothesis”
The aim of our study was to examine three rapid visual field tests HSF, TOP, and FDT and determine their diagnostic value to detect glaucomatous field defects in patients with glaucoma.
What type of study?

**Primary** – these report research first hand.

- *Experimental* –

- *Clinical trials* – intervention offered.

- *Observational* – something is measured in a group.
What type of study?

**Secondary** – summarise and draw conclusions from primary studies.

- **Overview**
  - Non systematic reviews (summary)
  - Systematic reviews (rigorous and pre-defined methodology)
  - Meta-analyses (integration of numerical data from more than one study)

- **Guidelines** (leads to advice on behaviour)
- **Decision analyses** (to help make choices for doctor or patient)
- **Economic analyses** (i.e. is this a good use of resources?)
The Hierarchy of Evidence

1. Systematic reviews & meta-analyses
2. Randomised controlled trials
3. Cohort studies
4. Case-control studies
5. Cross sectional surveys
6. Case reports
7. Expert opinion
8. Anecdotal
Specific types of study

- Did investigator assign exposures?
  - Yes: Experimental study
    - Random allocation?
      - Yes: Randomised controlled trial
      - No: Non-randomised controlled trial
  - No: Observational study
    - Comparison group?
      - Yes: Analytical study
      - No: Descriptive study

Exposure → Outcome

Exposure and outcome at the same time

- Cohort study
- Case-control study
- Cross-sectional study
Was the study design appropriate?

- Fields of research
  - *Therapy*: testing the efficacy of drug treatments, surgical procedures, alternative methods of service delivery, or other interventions. Preferred study design is randomized controlled trial
  - *Diagnosis*: demonstrating whether a new diagnostic test is valid (can we trust it?) and reliable (would we get the same results every time?). Preferred study design is cross sectional survey in which both the new test and the gold standard are performed
Was the study design appropriate?-2

- **Screening**: demonstrating the value of tests which can be applied to large populations and which pick up disease at a presymptomatic stage. Preferred study design is *cross sectional survey*

- **Prognosis**: determining what is likely to happen to someone whose disease is picked up at an early stage. Preferred study design is *longitudinal cohort study*

- **Causation**: determining whether a putative harmful agent, such as environmental pollution, is related to the development of illness. Preferred study design is *cohort or case-control study*, depending on how rare the disease is, but *case reports* may also provide crucial information
1. Check the Title

- Read the title and check that you understand its meaning. Sometimes titles are inaccurate and do not reflect the content of the paper which follows.

- For example, one title indicating the use of a drug in the treatment of hypertension, prefaced a paper which merely described a short haemodynamic study.
1. Check the Title

- Watch for cryptic titles. Sometimes a useful paper may be hidden behind an indifferent title.
- Never rely on the title alone to accept or reject a paper for more detailed reading.
2. Who are the Authors?

- Range of expertise: professional backgrounds with address
- Research center?
- Principle researcher
- Number of authors
- Have any of the authors obvious connections with the drug industry?
3. Read the abstract

- This is a synopsis of the paper, which should give the **objective** of the study, the **methods** used, the **results** obtained and the **conclusions** reached.
3. **Read the abstract**

Beware of the following **warning signs**:

- 1. Confusion and possible contradictory statements - a good abstract should be crystal clear.
- 2. Overuse of statistical terms (especially p values).
- 3. Disparity between the number of subjects mentioned in the summary and the number in the paper.
4. Check the Introduction

- Check that a brief review of available background literature is provided and that the question being asked in the study follows logically from the available evidence.
Introduction

- General, concise description of problem
  - background to the work
  - previous research
- Where that work is deficient
  - how your research will be better
- State the hypothesis
- About 3 to 4 paragraphs
Methods

- Study design
- Participants
- Ethical approval
- Sample size
- Questionnaires
- Interventions
- Clinical assessments
- Statistical methods
5. Assessing Methodology:
Six essential questions
Six essential questions

1. Was the study original?
2. Who is it about?
3. Was the design of the study sensible?
4. Was bias avoided?
5. Was assessment "blind"?
6. Were preliminary statistical questions dealt with?
Six essential questions:

1. Was the study original?
   - Is this study bigger, continued for longer, or otherwise more substantial than the previous one(s)?
   - Will the numerical results of this study add significantly to a meta-analysis of previous studies?
   - Is the population that was studied different in any way?
   - Is the clinical issue addressed of sufficient importance, and is there sufficient doubt in the minds of the public or key decision makers?
Six essential questions:

2. Who is it about?

- How recruited?
  - Recruitment bias
- Who included?
- Who excluded?
- Studied in “real life circumstances”?
Six essential questions:

3. Was the design of the study sensible?

☐ What specific intervention or manoeuvre was being considered and what was it being compared to?

☐ What outcome was measured and how?
Six essential questions:

4. Was bias avoided?

- i.e. was it adequately controlled for?


Cohorts – population differences

Case control – true diagnosis, recall (and influences)
Six essential questions:

5. Was assessment "blind"?

If I knew that a patient had been randomised to an active drug to lower blood pressure rather than to a placebo, I might be more likely to recheck a reading which was surprisingly high. This is an example of performance bias, a pitfall for the unblinded assessor.
Six essential questions:

6. Were preliminary statistical questions dealt with?

- Statistical tests
- The size of the study
  - “power”
- The duration of follow-up
- The completeness of follow-up
  - “drop-outs”
6. Results

What was found?

- Should be logical – simple complex
Cheat on statistical tests

- Throw all your data into a computer and report as significant any relation where P<0.05
- If baseline differences between the groups favour the intervention group, remember not to adjust for them
- Do not test your data to see if they are normally distributed. If you do, you might get stuck with non-parametric tests, which aren't as much fun
- Ignore all withdrawals (drop outs) and non-responders, so the analysis only concerns subjects who fully complied with treatment
Always assume that you can plot one set of data against another and calculate an "r value" (Pearson correlation coefficient), and assume that a "significant" r value proves causation.

- If outliers (points which lie a long way from the others on your graph) are messing up your calculations, just rub them out. But if outliers are helping your case, even if they seem to be spurious results, leave them in.

- If the confidence intervals of your result overlap zero difference between the groups, leave them out of your report. Better still, mention them briefly in the text but don't draw them in on the graph—and ignore them when drawing your conclusions.
If the difference between two groups becomes significant four and a half months into a six month trial, stop the trial and start writing up. Alternatively, if at six months the results are "nearly significant," extend the trial for another three weeks.

If your results prove uninteresting, ask the computer to go back and see if any particular subgroups behaved differently. You might find that your intervention worked after all in Chinese women aged 52-61.
Does the y-axis start at zero?

- The y-axis should always begin at zero. If this is not so, someone is trying to make you believe that one of the groups has reached the lowest rate or number possible when this is not the case.
Four possible outcomes from any study

1. Difference is clinically important and statistically significant i.e. important and real.
2. Of clinical importance but not statistically significant. sample size too small.
3. Statistically significant but not clinically important i.e. not clinically meaningful.
4. Neither clinically important nor statistically significant.
7. Discussion

- Check that the progress in argument to the conclusion is *logical* and also that any doubts or *inconsistencies* which have been raised in your mind by earlier parts of the paper, are dealt with.

- Are *limitations* mentioned?

- Authors’ *speculations* should be clearly distinguished from results, and should be seen as *opinion* not *fact*. 
8. Bibliography

- If you find statements in the paper which you consider to be important check that a reference is provided.

- Be suspicious if no reference is given, or if the references which are provided are dated, or predominantly in obscure journals.
9. Acknowledgment

- Who? (and what)?

- Source of funding? (conflict of interest)
Quality of reporting ≠ quality of study

- It may be necessary to contact the authors for further information about aspects of the study or to collect raw data.
Recommended Reading

Trisha Greenhalgh : How to read a paper; the basis of evidence based medicine

- Gordon Guyatt, Drummond Rennie. Users’ Guides To The Medical Literature, A Manual for Evidence-Based Clinical Practice
Medical Statistics

Dr. Gholamreza Khalili
Department of Epidemiology & Biostatistics
School of public health
Tehran University of Medical Sciences
Why does good evidence from research fail to get into practice?

- 75% cannot understand the statistics
- 70% cannot critically appraise a research paper

Using Research for Practice: A UK Experience of the barriers scale
Dunn V, Crichton C, Williams K, Roe B, Seers K
Cleansing your inner data
Pump yourself up!
100% confident!
Finding your me-value

Raw chicken soup for the outbreak investigator's soul
Decorating your 2x2 table
Rejecting the null, embracing the whole
Why is statistics necessary?

- 58% of the population had GERD
- Mean age of the respondents was 25±8
- 25% of women and 50% of men lied about their age (4 in each group!?)
- Doctors live longer than normal people
Why is statistics necessary?

- **Descriptive statistics**
  - 58% of the population had GERD
  - Mean age of the respondents was 25±8

- **Inferential statistics**
  - 25% of women and 50% of men lied about their age (4 in each group!?)
  - Doctors live longer than normal people.
Descriptive statistics

- Point estimates: Mean, median, mode, relative frequency
- Distribution: Standard deviation
Inferential statistics: exploring associations and differences
Differences

- Continuous variables (blood pressure, age): 109±11 vs. 140±10

- Categorical variables (proportion of blind people): 10% vs. 2%
# Measures of Association

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>dead</th>
<th>alive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Risk (RR)</td>
<td>Medical</td>
<td>404</td>
<td>921</td>
<td>1325</td>
</tr>
<tr>
<td>Odds Ratio (OR)</td>
<td>CABG</td>
<td>350</td>
<td>974</td>
<td>1324</td>
</tr>
<tr>
<td>Absolute risk reduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number needed to treat</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Number Needed to Treat (NNT)

- only defined for a specific intervention!
- only defined for a specific outcome!
  - eg. Pravastatin™ 40 mg nocte x10 years, in a 65 year old male, ex-smoker with high BP and Diabetes, to reduce MI or Death.
- NNT is the inverse of Absolute Risk Reduction: i.e. 
  
  \[ \text{NNT} = \frac{1}{\text{ARR}} \]
Measures of Association

- Linear Correlation
  - Conditions
  - $r$ and $p$, CI

- Regression
  - Univariate
  - Multiple Regression
  - Logistic Regression
  - Cox Proportional Hazard Model

- Do they mean causation?
Associations may be due to

- **Chance (random error)**
  - statistics are used to reduce it by appropriate design of the study
  - statistics are used to estimate the probability that the observed results are due to chance

- **Bias (Systematic error)**
  - must be considered in the design of the study

- **Confounding**
  - can be dealt with during both the design and the analysis of the study

- **True association**
Dealing with chance error

- During design of study
  - Sample size
  - Power

- During analysis (Statistical measures of chance)
  - Test of statistical significance (P value)
  - Confidence intervals
Statistical measures of chance I
(Test of statistical significance)

<table>
<thead>
<tr>
<th>Observed association</th>
<th>Association in Reality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>
The p-value in a nutshell

Could the result have occurred by chance?

The result is **unlikely** to be due to chance

The result is **likely** to be due to chance

\[ p < 0.05 \]
a statistically significant result

\[ p > 0.05 \]
not a statistically significant result

\[ p = 0.05 \]
\( \frac{1}{20} \) or 1 in 20
result fairly unlikely to be due to chance

\[ p = 0.5 \]
\( \frac{1}{2} \) or 1 in 2
result quite likely to be due to chance
Significantitis

- Significantitis is the plague of our time.
  - A. Etemadi, 21st century epidemiologist

- The drug reduced blood pressure by 1mmHg
  \( p < 0.0000000000000001 \)
Confidence Interval (CI)

Is the range within which the true size of effect (never exactly known) lies, with a given degree of assurance (usually 95%)
The ACE inhibitor group had a 5% (95% CI: 1-9) higher survival.
Associations may be due to

- **Chance (random error)**
  - statistics are used to reduce it by appropriate design of the study
  - statistics are used to estimate the probability that the observed results are due to chance

- **Bias (Systematic error)**
  - must be considered in the design of the study

- **Confounding**
  - can be dealt with during both the design and the analysis of the study

- **True association**
Types of Bias

- Selection bias – identification of individual subjects for inclusion in study on the basis of either exposure or disease status depends in some way on the other axis of interest

- Observation (information) bias – results from systematic differences in the way data on exposure or outcome are obtained from the various study groups
Associations may be due to

- **Chance (random error)**
  - statistics are used to reduce it by appropriate design of the study
  - statistics are used to estimate the probability that the observed results are due to chance

- **Bias (Systematic error)**
  - must be considered in the design of the study

- **Confounding**
  - can be dealt with during both the design and the analysis of the study

- **True association**
Confounding

- Coffee
  - Smoking
    - Pancreatic cancer
Confounding

- Coffee
- Smoking
- Pancreatic cancer
Confounding

Possible cause

confounder

effect
Associations may be due to

- **Chance (random error)**
  - statistics are used to reduce it by appropriate design of the study
  - statistics are used to estimate the probability that the observed results are due to chance

- **Bias (Systematic error)**
  - must be considered in the design of the study

- **Confounding**
  - can be dealt with during both the design and the analysis of the study

- **True association**
DETERMINATION OF CAUSATION

- The general QUESTION: Is there a cause and effect relationship between the presence of factor X and the development of disease Y?
Nature of Evidence:

1. Replication of Findings –
   - consistent in populations

2. Strength of Association –
   - significant high risk

3. Temporal Sequence –
   - exposure precede disease
Nature of Evidence:

4. Dose-Response –
   - higher dose exposure, higher risk

5. Biologic Credibility –
   - exposure linked to pathogenesis

6. Consideration of alternative explanations –
   - the extent to which other explanations have been considered.
Nature of Evidence

7. Cessation of exposure (Dynamics) –
   - removal of exposure – reduces risk

8. Specificity
   - specific exposure is associated with only one disease

9. Experimental evidence
What is the appropriate test?

- Scales
  - Nominal
  - Ordinal
  - Interval
  - Ratio
Normal Distribution

Figure 5–6. Standard normal (z) distribution.
Skewed curve
- Parametric versus non-parametric tests
- Transformation
Subgroup analysis?

Analyses showed that the drug was especially effective in women above 35 who were unable to say supercalifragilisticexpialidocious.

We divided the study population according to sex, then each group were divided to 10 age groups, each age group was subdivided according to educational background and whether they were left-handed or right-handed.
Scenario

6.1.1 Example: Signs of the Zodiac
(Effect of new dietary control regime.)

Data: 250 subjects chosen ‘randomly’. Weighed at start of week and again at end of week. Data in kg.

Results:

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SE Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight before</td>
<td>250</td>
<td>58.435</td>
<td>12.628</td>
<td>0.799</td>
</tr>
<tr>
<td>Weight after</td>
<td>250</td>
<td>58.309</td>
<td>12.636</td>
<td>0.799</td>
</tr>
<tr>
<td>Difference</td>
<td>250</td>
<td>0.126</td>
<td>1.081</td>
<td>0.068</td>
</tr>
</tbody>
</table>

So, average weight loss is 0.13kg (≈1/4 pound)
Confidence interval for mean weight loss is (−0.009, 0.260)kg.
Paired t-test for weight loss gives a t-statistic of 0.184, giving a p-value of 0.067 (using a two-sided test). (t=0.126/0.068)

Not quite significant at the 5% level!
Subgroup analysis

Mean weight loss by sign of the Zodiac

<table>
<thead>
<tr>
<th>Zodiac sign</th>
<th>n</th>
<th>mean weight loss</th>
<th>standard error of mean</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aquarius</td>
<td>26</td>
<td>0.313</td>
<td>0.217</td>
<td>1.44</td>
<td>0.161</td>
</tr>
<tr>
<td>Aries</td>
<td>15</td>
<td>0.543</td>
<td>0.205</td>
<td>2.65</td>
<td>0.019</td>
</tr>
<tr>
<td>Cancer</td>
<td>21</td>
<td>0.271</td>
<td>0.249</td>
<td>1.09</td>
<td>0.289</td>
</tr>
<tr>
<td>Capricorn</td>
<td>27</td>
<td>-0.191</td>
<td>0.222</td>
<td>-0.86</td>
<td>0.397</td>
</tr>
<tr>
<td>Gemini</td>
<td>18</td>
<td>0.068</td>
<td>0.266</td>
<td>0.26</td>
<td>0.801</td>
</tr>
<tr>
<td>Leo</td>
<td>22</td>
<td>0.194</td>
<td>0.234</td>
<td>0.83</td>
<td>0.416</td>
</tr>
<tr>
<td>Libra</td>
<td>26</td>
<td>0.108</td>
<td>0.217</td>
<td>0.50</td>
<td>0.623</td>
</tr>
<tr>
<td>Pisces</td>
<td>19</td>
<td>0.362</td>
<td>0.232</td>
<td>1.56</td>
<td>0.136</td>
</tr>
<tr>
<td>Sagittarius</td>
<td>12</td>
<td>0.403</td>
<td>0.294</td>
<td>1.37</td>
<td>0.197</td>
</tr>
<tr>
<td>Scorpio</td>
<td>20</td>
<td>0.030</td>
<td>0.274</td>
<td>0.11</td>
<td>0.248</td>
</tr>
<tr>
<td>Taurus</td>
<td>22</td>
<td>-0.315</td>
<td>0.183</td>
<td>-1.72</td>
<td>0.099</td>
</tr>
<tr>
<td>Virgo</td>
<td>22</td>
<td>0.044</td>
<td>0.238</td>
<td>0.18</td>
<td>0.955</td>
</tr>
</tbody>
</table>
# Statistical Tests

<table>
<thead>
<tr>
<th>Goal</th>
<th>Type of Data</th>
<th>Measurement (from Gaussian Population)</th>
<th>Rank, Score, or Measurement (from Non-Gaussian Population)</th>
<th>Binomial (Two Possible Outcomes)</th>
<th>Survival Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describe one group</td>
<td></td>
<td>Mean, SD</td>
<td>Median, interquartile range</td>
<td>Proportion</td>
<td>Kaplan Meier survival curve</td>
</tr>
<tr>
<td>Compare one group to a hypothetical value</td>
<td>One-sample <em>t</em> test</td>
<td>Wilcoxon test</td>
<td>Chi-square or Binomial test **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compare two unpaired groups</td>
<td>Unpaired <em>t</em> test</td>
<td>Mann-Whitney test</td>
<td>Fisher's test (chi-square for large samples)</td>
<td>Log-rank test or Mantel-Haenszel*</td>
<td></td>
</tr>
<tr>
<td>Compare two paired groups</td>
<td>Paired <em>t</em> test</td>
<td>Wilcoxon test</td>
<td>McNemar's test</td>
<td>Conditional proportional hazards regression*</td>
<td></td>
</tr>
</tbody>
</table>
**Statistical Tests**

<table>
<thead>
<tr>
<th><strong>Compare three or more unmatched groups</strong></th>
<th>One-way ANOVA</th>
<th>Kruskal-Wallis test</th>
<th>Chi-square test</th>
<th>Cox proportional hazard regression**</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compare three or more matched groups</strong></td>
<td>Repeated-measures ANOVA</td>
<td>Friedman test</td>
<td>Cochrane Q**</td>
<td>Conditional proportional hazards regression**</td>
</tr>
<tr>
<td><strong>Quantify association between two variables</strong></td>
<td>Pearson correlation</td>
<td>Spearman correlation</td>
<td>Contingency coefficients*</td>
<td>*</td>
</tr>
<tr>
<td><strong>Predict value from another measured variable</strong></td>
<td>Simple linear regression or Nonlinear regression</td>
<td>Nonparametric regression**</td>
<td>Simple logistic regression*</td>
<td>Cox proportional hazard regression*</td>
</tr>
<tr>
<td><strong>Predict value from several measured or binominal variables</strong></td>
<td>Multiple linear regression* or Multiple nonlinear regression**</td>
<td>Multiple logistic regression*</td>
<td>Cox proportional hazard regression*</td>
<td></td>
</tr>
</tbody>
</table>

*Please note that the table contains various statistical tests and their corresponding descriptions.** indicates advanced or specialized techniques.
CHECK-LISTS AND TOOLS

Dr. GR. Khalili
Department of Epidemiology & Biostatistics
School of public health
Tehran University of Medical Sciences
What is critical appraisal?

- Critical appraisal is the **assessment** of evidence by **systematically** reviewing its **relevance**, **validity** and **results** to specific situations.

<table>
<thead>
<tr>
<th>Critical appraisal is <strong>not:</strong></th>
<th>Critical appraisal is:</th>
</tr>
</thead>
<tbody>
<tr>
<td>× Negative dismissal of any piece of research</td>
<td>✓ Balanced assessment of benefits and strengths of research against its flaws and weaknesses</td>
</tr>
<tr>
<td>× Assessment of results alone</td>
<td>✓ Assessment of research process and results</td>
</tr>
<tr>
<td>× Based entirely on detailed statistical analysis</td>
<td>✓ Consideration of quantitative and qualitative aspects of research</td>
</tr>
<tr>
<td>× To be undertaken by expert researchers/statisticians only</td>
<td>✓ To be undertaken by all health professionals as part of their work</td>
</tr>
</tbody>
</table>
Key Steps To Effective Critical Appraisal

1. What are the results?
2. Are the Results valid?
3. How will these results help me/my colleagues do their job?
Critical Appraisal Tools

- Why do we need them?
- Where we can find them?
CASP

- The Critical Appraisal Skills Programme (CASP) was developed in Oxford in 1993 and has over the past years helped to develop an evidence based approach in health and social care.
- The CASP appraisal tools are based on the guides produced by the Evidence Based Medicine Working Group, a group of clinicians at McMaster university, Hamilton, Canada, and colleagues across North America, published in the Journal of the American Medical Association.
CASP ...

- Systematic Reviews
- Randomized Controlled Trials (RCTs)
- Qualitative Research
- Economic Evaluation Studies
- Cohort Studies
- Case Control Studies
- Diagnostic Test Studies
Three questions

☐ Valid?
Is the methodology appropriate to answer the question.
Is it carried out in a sound way, eliminating bias and confounding?

☐ Reliable?
Are the results real or because of chance?

☐ Applicable?
Will the results help locally?
Center for evidence based medicine

- Offers Critical Appraisal Sheets
International Centre for Allied Health Evidence

- Appraising randomized controlled trials
- Appraising non-randomized controlled trials
- Appraising other forms of quantitative research
- Appraising case studies
- Appraising qualitative research
- Appraising mixed methods research
- Appraising systematic reviews
- Appraising meta-analyses
- Appraising clinical guidelines
- Appraising outcome measures
- Assessing treatment choices
AGREE

- Appraisal of Guidelines Research and Evaluation
- The AGREE Instrument for the assessment of clinical practice guidelines is available online in several languages

http://www.agreecollaboration.org
DISCERN

Quality criteria for consumer health information on treatment choices

- **DISCERN** is a brief questionnaire which provides users with a valid and reliable way of assessing the quality of written information on treatment choices for a health problem.

- DISCERN can also be used by authors and publishers of information on treatment choices as a guide to the standard which users are entitled to expect.

- New **DISCERN Genetics** site provides a reliable way of assessing the quality of information on genetic testing and screening.
STROBE Statement

- STrengthening the Reporting of Observational studies in Epidemiology

- www.strobe-statement.org
Some other fields …

- Economic Evaluation Studies
- Diagnostic Test Studies
- Qualitative Studies
- Therapy
- Diagnosis
- Harm
- Prognosis
- Quantitative Research Studies
- Interventions Addressing the Need for education
- Library Research
- meta-analyses
Do they work?

- Katrak et al. systematic review of the content of critical appraisal tools. *BMC Medical Research Methodology* 2004

- Few critical appraisal tools had documented evidence of validity of their items, or reliability of use.
Appraisal Tools for

OBSERVATIONAL STUDIES
Types of Observational studies

- Cohort
- Case-control
- Cross-sectional
- Ecologic
- Case series
- Case report
STROBE Statement

- STrengthening the Reporting of OBservational studies in Epidemiology
- Many journals refer to the STROBE Statement in their Instructions for Authors.
- Provides recommendation for each section (22 items)
Available STROBE check-lists

- STROBE checklist for cohort, case-control, and cross-sectional studies (combined)
- Checklist for cohort studies
- Checklist for case-control studies
- Checklist for cross-sectional studies
Title and abstract

(a) Indicate the study’s design with a commonly used term in the title or the abstract

(b) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction

- **Background/rationale:**
  - Explain the scientific background and rationale for the investigation being reported

- **Objectives:**
  - State specific objectives, including any pre-specified hypotheses
Methods

- **Study design**
  - Present key elements of study design early in the paper

- **Setting**
  - Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Methods: participants

- **Cohort study**—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up
- **Case-control study**—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls
- **Cross-sectional study**—Give the eligibility criteria, and the sources and methods of selection of participants
Methods: matched studies

- **Cohort study**—For matched studies, give matching criteria and number of exposed and unexposed

- **Case-control study**—For matched studies, give matching criteria and the number of controls per case
Methods: Variables

- Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.

- Quantitative variables
  - Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why.
Methods: Data sources/ measurement

- For each variable of interest, give sources of data and details of methods of assessment (measurement).

- Describe comparability of assessment methods if there is more than one group.
Method: Bias & Study size

- Describe any efforts to address potential sources of bias
- Explain how the study size was arrived at
Method : Statistical methods

- Describe all statistical methods, including those used to control for confounding
- Describe any methods used to examine subgroups and interactions
- Explain how missing data were addressed
- Cohort study—If applicable, explain how loss to follow-up was addressed
- Case-control study—If applicable, explain how matching of cases and controls was addressed
- Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy
- Describe any sensitivity analyses
Results: Participants

- Report numbers of individuals at each stage of study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed.

- Give reasons for non-participation at each stage.

- Consider use of a flow diagram.
Results: Descriptive data

- characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders
- number of participants with missing data for each variable of interest
- *Cohort study*—Summarise follow-up time (e.g., average and total amount)
Results: Outcome data

- **Cohort study**—Report numbers of outcome events or summary measures over time
- **Case-control study**—Report numbers in each exposure category, or summary measures of exposure
- **Cross-sectional study**—Report numbers of outcome events or summary measures
Main results and Other analyses

- unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included.

- *Report category boundaries when continuous variables were categorized.*

- If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period.

- Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses.
Discussion

- **Key results**: Summarize key results with reference to study objectives
- **Limitations**: Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
- **Interpretation**: Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
- **Generalisability**: Discuss the generalisability (external validity) of the study results
Other information

- the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based
CASP: Cohort study

CRITICAL APPRAISAL SKILLS PROGRAMME
making sense of evidence

12 questions to help you make sense of a cohort study

Public Health Resource Unit, Oxford
General comments

☐ Three broad issues need to be considered when appraising a cohort study:

- *Are the results of the study valid?*
- *What are the results?*
- *Will the results help locally?*
screening questions

The first two questions are screening questions and can be answered quickly. If the answer to those two is "yes", it is worth proceeding with the remaining questions.

1. Did the study address a clearly focused issue?

2. Did the authors use an appropriate method to answer their question?
Validity: Detailed Questions

3. Was the cohort recruited in an acceptable way?

4. Was the exposure accurately measured to minimize bias?

5. Was the outcome accurately measured to minimize bias?

6. Have the authors identified all important confounding factors? Have they taken account of the confounding factors in the design and/or analysis?

7. Was the follow up of subjects complete enough? Was the follow up of subjects long enough?
What are the results?

8. What are the results of this study?

9. How precise are the results? How precise is the estimate of the risk?

10. Do you believe the results?
11. Can the results be applied to the local population?

12. Do the results of this study fit with other available evidence?
Appraisal Tools for
RANDOMIZED
CONTROLLED TRIALS
CONSORT

- Consolidated Standards of Reporting Trials
- 22 items
HISTORY OF CONSORT

- CONSORT (Consolidated Standards of Reporting Trials) statement (In the mid 1990s)
- The revised CONSORT statement (1999, 2000)
The CONSORT statement comprises:

- a 22-item checklist pertaining to the content of the Title, Abstract, Introduction, Methods, Results, Comment.
- a flow diagram depicting information from 4 stages of a trial enrollment, intervention allocation, follow-up, analysis.
Title and abstract

- How participants were allocated to interventions (e.g., “random allocation,” “randomized,” or “randomly assigned”).
Introduction: Background

- Scientific background and explanation of rationale.
Method:

- **Participants**: Eligibility criteria for participants and the settings and locations where the data were collected.

- **Interventions**: Precise details of the interventions intended for each group and how and when they were actually administered.

- **Objectives**: Specific objectives and hypotheses.
Method:

☐ Outcomes: Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors).

☐ Sample size: How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.
Method: Randomization

- **Sequence generation:** Method used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification).
- **Allocation concealment:** Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.
- **Implementation:** Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.
Method:

- **Blinding** (masking): Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. If done, how the success of blinding was evaluated.

- **Statistical methods**: Statistical methods used to compare groups for primary outcome(s); methods for additional analyses, such as subgroup analyses and adjusted analyses.
Results

- **Participant flow:** Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.
- **Recruitment:** Dates defining the periods of recruitment and follow-up.
- **Baseline data:** Baseline demographic and clinical characteristics of each group.
Results

- **Numbers analyzed:** Number of participants (denominator) in each group included in each analysis and whether the analysis was by “intention to treat.” State the results in absolute numbers when feasible (e.g., 10 of 20, not 50%).

- **Outcomes and estimation:** For each primary and secondary outcome, a summary of results for each group and the estimated effect size and its precision (e.g., 95% confidence interval).
Results

- **Ancillary analyses:** Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory.

- **Adverse events:** All important adverse events or side effects in each intervention group
Discussion

- **Interpretation**: Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes.

- **Generalizability**: Generalizability (external validity) of the trial findings.

- **Overall evidence**: General interpretation of the results in the context of current evidence.
Assessed for eligibility (n = ...)

Excluded (n = ...)
- Did not meet inclusion criteria (n = ...)
- Refused to participate (n = ...)
- Other reasons (n = ...)

Randomized (n = ...)

Allocated to intervention (n = ...)
- Received allocated intervention (n = ...)
- Did not receive allocated intervention (give reasons) (n = ...)

Allocated to intervention (n = ...)
- Received allocated intervention (n = ...)
- Did not receive allocated intervention (give reasons) (n = ...)

Lost to follow-up (n = ...)
(give reasons)

Discontinued intervention (give reasons) (n = ...)

Lost to follow-up (n = ...)
(give reasons)

Discontinued intervention (give reasons) (n = ...)

Analysed (n = ...)
- Excluded from analysis (give reasons) (n = ...)

Analysed (n = ...)
- Excluded from analysis (give reasons) (n = ...)

Enrollment
CASP

Critical Appraisal Skills Programme (CASP)
making sense of evidence

10 questions to help you make sense of randomised controlled trials

Screening Questions

1. Did the study ask a clearly-focused question?

2. Was this a randomised controlled trial and was it appropriately so?
Detailed Questions

3. Were participants appropriately allocated to intervention and control groups?
4. Were participants, staff and study personnel ‘blind’ to participants’ study group?
5. Were all of the participants who entered the trial accounted for at its conclusion?
6. Were the participants in all groups followed up and data collected in the same way?
7. Did the study have enough participants to minimize the play of chance?
8. How are the results presented and what is the main result?
9. How precise are these results?
10. Were all important outcomes considered so the results can be applied?
Appraisal Tools for

DIAGNOSTIC TESTS
خصوصیات ذاتی یک تست

- حساسیت: احتمال تشخیص صحیح آزمون در افراد واقعا بیمار

- ویژگی: احتمال تشخیص صحیح آزمون در افراد واقعا سالم

<table>
<thead>
<tr>
<th>test</th>
<th>Gold standard</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>TN + FP</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>FP</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>FN</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>TN</td>
<td></td>
</tr>
</tbody>
</table>


\[
\text{Sensitivity} = \frac{TP}{TP + FN} \\
\text{Specificity} = \frac{TN}{TN + FP}
\]
 محاسبه ویژگی و حساسیت

<table>
<thead>
<tr>
<th>Gold Standard</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>8</td>
</tr>
<tr>
<td>-</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
</tr>
</tbody>
</table>

Sensitivity = 40%
Specificity = 90%
Gold Standard

\[
\begin{array}{cc}
+ & - \\
8 & 2 \\
12 & 18 \\
\hline
20 & 20 \\
\end{array}
\]

Sensitivity = 40%
Specificity = 90%

Gold Standard

\[
\begin{array}{cc}
+ & - \\
15 & 4 \\
5 & 16 \\
\hline
20 & 20 \\
\end{array}
\]

Sensitivity = 75%
Specificity = 80%
سایر خصوصیات تست

ارزش اخباری مثبت: نسبت افراد واقعا بیمار در بین کسانی که نتیجه تست آنان مثبت است.

ارزش اخباری منفی: نسبت افراد واقعا سالم در بین کسانی که نتیجه تست آنان منفی است.

### Gold standard

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>TP</td>
<td>FP</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>FN</td>
<td>TN</td>
<td></td>
</tr>
</tbody>
</table>

### PPV

\[
PPV = \frac{TP}{TP + FP}
\]

### NPV

\[
NPV = \frac{TN}{TN + FN}
\]
محاسبه ارزش اخباری

PPV = 80%
NPV = 60%

<table>
<thead>
<tr>
<th>Test</th>
<th>Gold Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>8</td>
</tr>
<tr>
<td>-</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>18</th>
<th>30</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>2</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

اداره انتشارات و علم سنجی دانشگاه علوم پزشکی تهران
تاثیر شیوع بر ارزش اخباری

<table>
<thead>
<tr>
<th></th>
<th>Gold Standard</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>-</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>40</td>
</tr>
</tbody>
</table>

PPV = 80%
NPV = 60%

<table>
<thead>
<tr>
<th></th>
<th>Gold Standard</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>-</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>40</td>
</tr>
</tbody>
</table>

PPV = 40%
NPV = 73%
تغییرات ارزش احتمال مثبت بر حسب شیوع با تابت
ماندن حساسیت و ویژگی

<table>
<thead>
<tr>
<th>prevalence</th>
<th>PPV</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>1.8</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>1</td>
<td>15.4</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>5</td>
<td>48.6</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>50</td>
<td>94.7</td>
<td>90</td>
<td>95</td>
</tr>
</tbody>
</table>
Reliability

- Reliability refers to the *repeatability* or reproducibility of a test.

- It can be assessed by *repeating the test* using the same or different observers.
Validity

- Relates to whether the test measures what it purports to measure. Is the result true?

- It can be assessed by comparing the test results with a Gold Standard.
Likelihood ratio

LR Positive = \frac{\text{Likelihood of (+) test in diseased persons}}{\text{Likelihood of (+) test in healthy persons}}

LR Positive = \frac{\text{Sensitivity}}{1 - \text{Specificity}}

LR Negative = \frac{\text{Likelihood of (-) test in diseased persons}}{\text{Likelihood of (-) test in healthy persons}}

LR Negative = \frac{1 - \text{Sensitivity}}{\text{Specificity}}
Continuous Measurements

Cutoff Value for Positive Test

Proportion

Healthy

Diseased

TN

TP

FN

FP

SBP
Continuous Measurements

Cutoff Value for Positive Test

Healthy

Diseased

Proportion

TN

FP

FN

TP

SBP
### Continuous Measurements

<table>
<thead>
<tr>
<th>Healthy</th>
<th>Diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN</td>
<td>TP</td>
</tr>
<tr>
<td>FN</td>
<td>FP</td>
</tr>
</tbody>
</table>

#### Cutoff Value for Positive Test

- **Proportion**
- **Healthy**
- **Diseased**
- **TN**
- **TP**
- **FN**
- **FP**
- **SBP**
Receiver operator curves

- By plotting the sensitivity and specificity of a test for different cut off points a ROC can be produced which helps illustrate the optimum cut off point to use.
Receiver Operator Characteristic Curve
ROC Curve

Sensitivity

1 - Specificity
Standards for Reporting of Diagnostic Accuracy (STARD)

Improve the *accuracy* and completeness of research reporting and allow readers to assess the “potential for bias” in the study reported.

Always use:
- FLOW CHART or Diagram
- CHECKLIST
<table>
<thead>
<tr>
<th>Section &amp; Topic</th>
<th>Item #</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE/ABS.../KEYWORDS</td>
<td>1</td>
<td>Identify the article as a study of diagnostic accuracy (recommend <a href="https://www.ncbi.nlm.nih.gov/mesh">MeSH</a> Heading 'sensitivity and specificity').</td>
</tr>
<tr>
<td>INTRODUCTION...</td>
<td>2</td>
<td>State the <strong>research questions</strong> or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.</td>
</tr>
</tbody>
</table>
METHODS

Participants
3 Describe the study population: The *inclusion and exclusion criteria*, setting and locations where the data were collected.

4 Describe **participant recruitment**: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?

5 **Describe participant sampling**: *Was the study population a consecutive series of participants* defined by the selection criteria in items 3 and 4? If not, specify how participants were further selected.

6 Describe data collection: Was data collection planned before the index test and reference standard were performed *(prospective study)* or after *(retrospective study)*?

Test methods
7 Describe the **reference standard** and its rationale.

8 Describe technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.

9 Describe definition of and rationale for the units, cutoffs and/or categories of the results of the index tests and the reference standard.

10 Describe the number, training and expertise of the persons executing and reading the index tests and the reference standard.

11 Describe whether or not the readers of the index tests and **reference standard were blind (masked) to the results of the other tests** and describe any other clinical information available to the readers.
# METHODS

<table>
<thead>
<tr>
<th>Section &amp; Topic</th>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical methods</td>
<td>12</td>
<td>Describe methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g., 95% confidence intervals).</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Describe methods for calculating test reproducibility, if done.</td>
</tr>
<tr>
<td>Section &amp; Topic</td>
<td>Item</td>
<td>#</td>
</tr>
<tr>
<td>----------------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>RESULTS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>14</td>
<td>Report when study was done, including beginning and ending dates of recruitment.</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Report <em>clinical and demographic characteristics of the study population</em> (e.g. age, sex, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centers).</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>Report the number of participants satisfying the criteria for inclusion that did or did not undergo the index tests and/or the reference standard; describe why participants failed to receive either test (a flow diagram is strongly recommended).</td>
</tr>
<tr>
<td>Test results</td>
<td>17</td>
<td>Report time interval from the index tests to the reference standard, and any treatment administered between.</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>Report distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>Report a cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>Report any adverse events from performing the index tests or the reference standard.</td>
</tr>
</tbody>
</table>
Estimates

21 Report estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).

22 Report how indeterminate results, missing responses and outliers of the index tests were handled.

23 Report estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.

24 Report estimates of test reproducibility, if done.
Discuss the clinical applicability of the study findings.
CASP

Critical Appraisal Skills Programme (CASP)
making sense of evidence

12 questions to help you make sense of a diagnostic test study

Three broad issues

- Are the results of the study valid?
- What are the results?
- Will the results help me and my patients/population?
Screening Questions

- **Was there a clear question for the study to address?**
  
  *A question should include information about:*
  
  - the population
  - the test
  - the setting
  - the outcomes

- **Was there a comparison with an appropriate reference standard?**
  
  *Is this reference test(s) the best available indicator in the circumstances?*
Detailed Questions

3. Did all patients get the diagnostic test and the reference standard?

4. Could the results of the test of have been influenced by the results of the reference standard?

5. Is the disease status of the tested population clearly described?

6. Were the methods for performing the test described in sufficient detail?
what are the results?

7. What are the results?

8. How sure are we about these results?
Will the results help me and my patients/population?

Consider whether you are primarily interested in the impact on a population or individual level.

9. Can the results be applied to your patients the population of interest?

10. Can the test be applied to your patient or population of interest?

11. Were all outcomes important to the individual or population considered?

12. What would be the impact of using this test on your patients/population?
Critical appraisal of

SECONDARY STUDIES
secondary study

- A secondary study does not generate any data from direct measurements, instead, it analyses a set of primary studies and usually seeks to aggregate the results from these in order to provide stronger forms of evidence about a particular phenomenon.
What is a systematic review?

- A review that has been prepared using some kind of systematic approach to minimising biases and random errors, and that the components of the approach will be documented in a materials and methods section.

Chalmers et al, 1995
What is a meta-analysis?

A statistical analysis of the results from independent studies, which generally aims to produce a single estimate of the treatment effect.

Egger et al, 2001
What is a systematic review

Reviews

Systematic reviews
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
Some of the Appraising tools

Appraising systematic reviews

- Critical Appraisal Skills Program (CASP): Systematic Reviews
- Systematic Review (of therapy) Worksheet
- ARIF (Aggressive Research Intelligence Facility)

Appraising meta-analyses

- QUOROM Statement Checklist
CASP

Critical Appraisal Skills Programme (CASP)
making sense of evidence

10 questions to help you make sense of reviews

Screening Questions

1. Did the review ask a clearly-focused question?

2. Did the review include the right type of study?
   – address the review’s question
   – have an appropriate study design
Detailed Questions

- 3. Did the reviewers try to identify all relevant studies?
  - which bibliographic databases were used
  - if there was personal contact with experts
  - if the reviewers searched for unpublished studies
  - if the reviewers searched for non-English-language studies
4. Did the reviewers assess the quality of the included studies?

- *if a clear, pre-determined strategy was used to determine which studies were included.*
  
  - *a scoring system*
  
  - *more than one assessor*
5. If the results of the studies have been combined, was it reasonable to do so?

– the results of each study are clearly displayed
– the results were similar from study to study
  (look for tests of heterogeneity)
– the reasons for any variations in results are discussed
Detailed Questions

6. How are the results presented and what is the main result?

– how the results are expressed (e.g. odds ratio, relative risk, etc.)
– how large this size of result is and how meaningful it is
– how you would sum up the bottom-line result of the review in one sentence
Detailed Questions

7. How precise are these results?

8. Can the results be applied to the local

9. Were all important outcomes considered?
   (individual, policy makers and professionals,
   family/caregivers, wider community)

10. Should policy or practice change as a result
    of the evidence contained in this review?
    (whether any benefit reported outweighs any harm and/or
    cost. If this information is not reported can it be filled in from
    elsewhere?)
Plagiarism

Dr. GR. Khalili
Department of Epidemiology & Biostatistics
School of public health
Tehran University of Medical Sciences
Authorship

- Authors are listed to provide a public record of responsibility and credit for the work.

- Only those who can take both responsibility and credit for a work should be authors.
Authorship

1. Substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data **AND**

2. Drafting the article or revising it critically for important intellectual content **AND**

3. Final approval of the version to be published

All 3 conditions should be met for assigning authorship

(Consensus Statement, International Council of Medical Journal Editors)
What is Plagiarism?

According to Webster's New World Dictionary, to plagiarize is to "take the ideas, writings, etc. from another and pass them off as one's own."

Plagiarism is the act of presenting the words, ideas, images, sounds, or the creative expression of others as your own.
سرقت علمی (Plagiarism)

- استفاده بدون ذکر مأخذ از ایده‌های اثر انتشار یافته‌یا انتشار نیافته، دیگران.
- تمام منابع باید ذکر شوند.
- اگر میزان قابل توجهی از اثر دیگری قرار است مورد استفاده قرار گیرد باید از وی اجازه گرفته شود.
Fraud

- در مقاله نویسی، Fraud درجه مختلفی دارد که بدترين درجه آن Fabrification است به این معنی که داده‌ها و Data سنتز گردید. را به این معنی که داده‌ها و اصلا وجود خارجی نداشته‌اند.
- يا به اشتباه اندامختن می‌باشد، Fraud درجه خفیف‌تر یعنی این‌که تعدادی از داده‌ها را حذف کنند و یا تعدادی را اضافه کنند و به طور کلی دستکاری داده‌ها را Falsification می‌گویند و یا این‌که تعدادی از داده‌ها را حذف کرده و کلی دستکاری داده‌ها را Falsification انجام دهند.
- اگر داده‌های عوارض جانبی را در یک مطالعه گزارش نکنید، باز هم به نوعی Falsification اتفاق افتاده است. clinical trial
How to avoid Plagiarism?

1. Quoting
2. Re-wording or Re-phrasing
Direct Quotes

- If you use someone else’s writing without putting it in quotes, you have blatantly plagiarized.
- Even if you add the source in your bibliography, it is still plagiarism!
Paraphrasing

- Be careful about rewriting someone else’s words. If your sentences use many of the same words and grammatical structure as the original source, it could be construed as plagiarism. Just put the text in your own words.
When Paraphrasing...

- Be sure you are not just rearranging or replacing words.
- Rewrite the phrase in your own words and credit the original source.
- Double check what you have wrote by comparing it with the original writing.
Let’s Practice: Are the Following Plagiarism or Not?

- **Original source:** The effort required to provide online information literacy instruction is intense.

- **Your paper:** ”The effort required to provide online information literacy instruction is intense.” (Smith 2006, p.42)
Answer?

Not Plagiarism

You’ve used quotation marks and cited the source so that no one believes that these are your own words.
Let’s Practice: Are the Following Plagiarism or Not?

**Original source:** The effort required to provide online information literacy instruction is intense.

**Your paper:** Smith (2006, p.42) argues that providing online courses in information literacy is hard work.
Answer?

Not Plagiarism

You’ve interpreted without quoting, have cited your source, and have not used a great deal of the original terminology.
Let’s Practice: Are the Following Plagiarism or Not?

**Original source:** The effort required to provide online information literacy instruction is intense.

**Your paper:** The work needed to provide online information literacy teaching is intense.
Plagiarism

Your version is a paraphrase of the original with a lot of the original terminology still there as well as the same sentence structure.
There are lots of Plagiarism Detection Softwares which you can find on the web.

Here are some of them:

- http://www.anticutandpaste.com
- http://www.plagiarismdetect.com
- http://www.dustball.com/cs/plagiarism.checker/
- http://www.millikin.edu/wcenter/plagiarism3.html
There are lots of Plagiarism Detection Softwares which you can find on the web. Here are some of them:

- http://www.anticutandpaste.com
- http://www.plagiarismdetect.com
- http://www.dustball.com/cs/plagiarism.checker/
- http://www.millikin.edu/wcenter/plagiarism3.html
Plagiarism Detection Websites

- There are lots of Plagiarism Detection Online Websites which you can find on the web.

- Here are some of them:
  - http://www.dupliplecker.com/
  - http://invention.swmed.edu/etblast/etblast.shtml
Dupli checker

8th grade math worksheets
Pre-algebra, exponents, geometry — Practice 200+ eighth-grade skills

Your University Assistant
We help U with Papers, Dissertation Online Courses for an A grade.

Search Options: Advanced Classic (New Feature)

Note: For Quicker Processing In Advanced Mode We Recommend Mozilla Firefox.

Please Enter Your Text Below And Press Search:

Note: You can also select content file (*.txt) by pressing Browse button & then press Insert Content Button.
An example

Sentence1: "take the ideas, writings, etc"
"take the ideas, writings, etc"....

Sentence2: "from another and pass them off as one's own"
http://www.stvincentscollege.edu/aboutsvc/whatisplagiarism.cfm
"from another and pass them off as one's own"....

http://www.yourdictionary.com/plagiarize
"from another and pass them off as one's own"....

http://www.lakesitepublications.com/tools5.html
"from another and pass them off as one's own"....

http://www.smccd.net/accounts/jackson/firstclass/course.html
"from another and pass them off as one's own"....
Guidelines for Publication

Widely cited guidelines for publication written by the International Committee of Medical Journal Editors. Guidelines can be found at

www.icmje.org
Guidelines for Publication

☐ Do not plagiarize!

- Always credit the work of others
- Be sure to cite sources
- Include all cited sources in the reference list and vice versa
- Obtain permission to include figures, models, graphs, etc.
Guidelines for Publication

- Adequately and accurately cite literature
  - Include adequate references to document ideas
  - Verify that referenced works are consistent with the ideas and information credited to them
  - Cite original sources whenever possible
  - Check the accuracy of citations so readers can locate referenced work
Publication Ethics Codes & Protocols

- You can find more guidelines & protocols of publication ethics in COPE (Committee of Publication Ethics) Website.

- http://www.publicationethics.org.uk/
Guidelines for Publication

- State in the manuscript when research has been approved by institutional review committees
- Acknowledge/disclose in the manuscript any real or perceived conflicts of interest to avoid the appearance of any bias
CONFLICT OF INTEREST
What is conflict of interest?

- Conflict of interest is a set of conditions in which professional judgement concerning a primary interest (such as patients' welfare or the validity of research) tends to be unduly influenced by a secondary interest (such as financial gain).

What is conflict of interest?

- Conflict of interest is a condition not a behaviour.
- Having a conflict of interest is not, in and of itself, evidence of wrongdoing.
- For many professionals, it is virtually impossible to avoid conflicts of interest from time to time.
- Reviewers?!
Potential Conflict of Interest

Possibility from the perspective of an independent observer that an individual’s private financial interest or family’s interests may influence professional actions, decisions, or judgment

- Not possible or desirable to eliminate
- Need to manage
Do you have a conflict of interest?

1. Have you in the past five years accepted the following from an organisation that may in any way gain or lose financially from the results of your study or the conclusions of your review, editorial, or letter:

☐ ____ Reimbursement for attending a symposium?
☐ ____ A fee for speaking?
☐ ____ A fee for organising education?
☐ ____ Funds for research?
☐ ____ Funds for a member of staff?
☐ ____ Fees for consulting?
Do you have a conflict of interest?

2. Have been employed by an organisation that may in any way gain or lose financially from the results of your study or the conclusions of your review, editorial, or letter?

3. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the results of your study or the conclusions of your review, editorial, or letter?

4. Do you have any other competing financial interests?
What should we do?

- In case of conflicting interests, one should declare.
- you might want to disclose any sort of competing interest that would embarrass you if it became generally known after publication.
Why don’t authors declare conflicts of interest?

- Some journals don’t require disclosure
- The culture is one of not disclosing
- Authors think that it’s somehow “naughty”
- Authors are confident that they are not affected by conflicts of interest

- What about reviewers?!
Does conflict of interest matter?

- Financial benefit makes doctors more likely to refer patients for tests, operations, or hospital admission, or to ask that drugs be stocked by a hospital pharmacy.

- Original papers published in journal supplements sponsored by pharmaceutical companies are inferior to those published in the parent journal.

- Reviews that acknowledge sponsorship by the pharmaceutical or tobacco industry are more likely to draw conclusions that are favourable to the industry.
Conflict of interest within journals

- Drug company sponsored supplements have been shown to be of inferior quality--but many journals publish them. They are a major source of income.
- Some journals exist simply to publish studies funded by pharmaceutical companies.
- Many journals depend heavily on advertising: does this influence their decisions on what to publish?
Conflict of interest within journals

- Some journals publish advertising next to related articles? Does this influence what they publish?
- Some journals make millions of dollars from reprints of articles--mostly of randomised trials funded by pharmaceutical companies
Conflict of interest within journals

- Acceptance of a particular study may be accompanied by a reprint order of more than a million dollars. It’s not difficult to tell which studies might produce such an order. Does this influence the decision on which studies to publish?

- Few (if any) journals publish the competing interests of their editors, editorial board, and management team and board
How to respond to conflict of interest?

- “If in doubt, disclose.”
- Sometimes the conflict will be so strong that it will forbid participation
- The danger of trying to eradicate conflict of interest is that it may encourage deception
Conclusions

- Concern about conflict of interest is not just political correctness
- Conflict of interest has an important impact on the information reaching health professionals and the public and on patient care
- Conflict of interest is very common in medicine