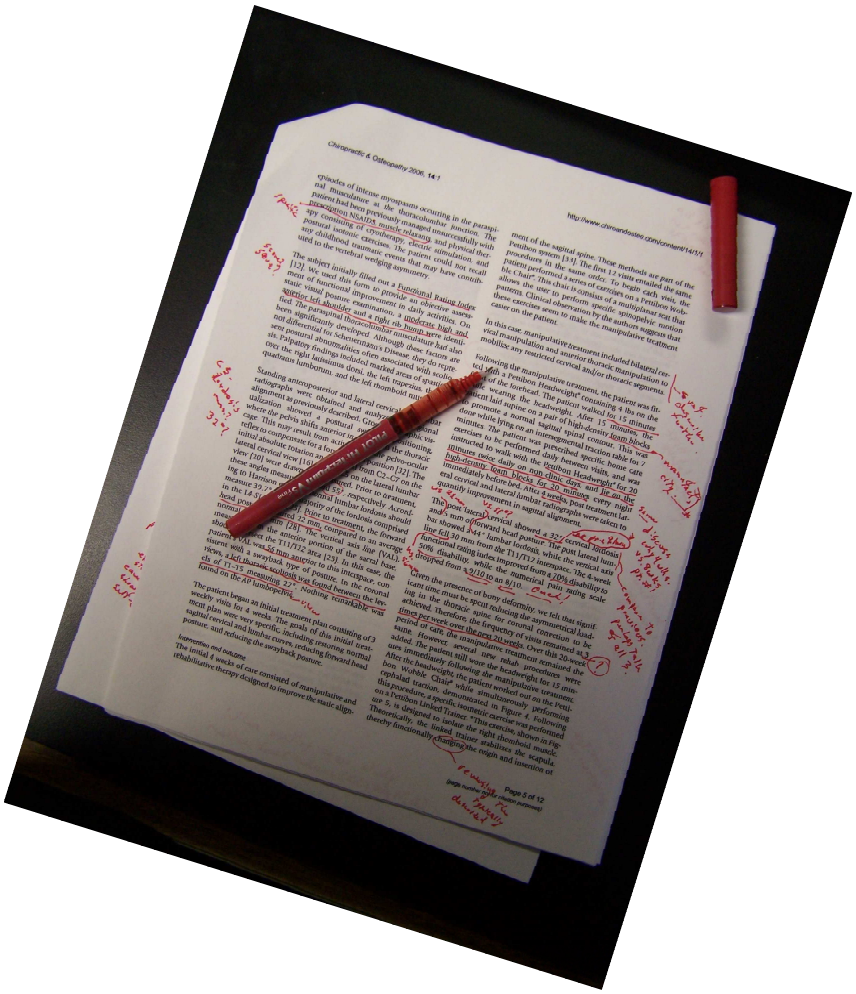
Critical Appraisal for Research Papers

Appraisal Checklist & Guide Questions



Critical Appraisal for Research Papers
Appraisal Checklist & Guide Questions

Name:

Title Topic:

PICO Question:

Bibliographic citation: (Vancouver format)

Search strategy review

Search Engines / Program(s):

Databases searched:
(indicate search engine if appropriate e.g., PubMed/MEDLINE or EBSCOhost/CINEHAL)

Key Search Terms:

Operators used:

Limits Used:

Additional Strategies used: (MeSH term, controlled vocabulary, clinical query options, etc.)

Search results:

Selection rationale:

Full text access:

Static link (citation / abstract URL):

Appraisal Checklist & Guide Questions

1) Study objectives and hypothesis

a) Purpose, objectives and hypothesis
   i) Using your words, what was the research question and objective(s) of the study?
   ii) Was the purpose of the study conveyed plainly and rationally?
   iii) Were the objectives of the study clearly stated?

2) Study design / evidence level

a) Study design and evidence level (CEBM hierarchy / strength):
   i) Was the study design stated and adequately described? What is the stated study design?
   ii) Considering the strengths and limitations of the study design, is it suitable for the objectives?
   iii) Was the study of adequate length in experimentation / observation / trial and measurement?
   iv) Were an adequate number of visits provided at appropriate intervals and frequency?

b) Before the Critical Appraisal, I consider this study (very strong, strong, moderate, weak, very weak)::

3) Methods: Subjects / Participants / Patient / Population

a) Population / Subjects / Sample:
   - In general, populations are large groups of people in a defined, described setting (urban, Metropolitan Chicago vs. rural India) or with a specific, defined characteristic (male between 28 and 45 years old, attendance at a specific hospital or private clinic).
   - A sample is a population subset which is selected from the larger population.

b) Selection / Inclusion & exclusion criteria / recruitment study site and circumstances
   - Selection bias refers to differences in the comparison groups attributable to incomplete randomization or specific allocation to a group. Patients may differ in additional factors or characteristics not under study that may affect the outcome of the disease (confounders). These factors (race, gender, age, nutritional status, socioeconomic status, etc.) may influence the outcomes or results of the study (reaction to an intervention or tendency toward a specific outcome such as illness).

     Selection bias is an issue when patients are chosen for an observation (or intervention), and must be dealt with in the design of a study (randomization or intentional allocation to a specific group. If confounding or multiple factors are possible, the entire population should either have or not have the factor, the factor should be present equally in all groups, or the results should be analyzed considering the presence and absence of the factor.

     Be aware of selection bias possibilities when groups are designated and defined as in non-randomized controlled clinical trials or study designs where subjects are not randomized into a specific group. Examine baseline differences between intervention and control groups or comparison groups and consider whether there are differences that might influence the outcome on their own, or contribute to a specific outcome.

     In RCTs, selection bias is avoided (theoretically) by blind, randomized allocation to a group without regard to any characteristics beyond any eligibility, inclusion or exclusion criteria. In observational studies such as cohort and case-control designs, selection of a comparable control group is difficult. Often, in cohort studies statistical adjustments are made for key variables in order to “control” differences contributing to selection bias (race, gender, age, etc.). In case-control studies (studies of very similar individuals with and without a particular disease or issue and analyzed retrospectively to determine cause or association), selection bias can be factor in defining whether the subject is a “case” or not. “Caseness” or classification of a subject as a case should be objectively defined and supported by outside references or guidelines. See also: randomization.
   i) Who was included in the study? Who was excluded from the study?
   ii) How were the subjects recruited?
   iii) Was there any selection bias?
   iv) “Where” did the study take place? Were the subjects studied in "real life" circumstances?

c) Base-line differences
   i) Did the population, experimental and control or comparison groups start with the same baseline demographics and prognostic factors? Did they have the same prognosis throughout the study?
   ii) Were there any differences noted between groups at the beginning of the trial?
   iii) Should any differences be considered as possible confounders?
   iv) If differences are present, are they acknowledged and addressed?
   v) Are any differences discussed as limitations?

d) Sample / Population size: what is / are the sample sizes, number of participants? Was the sample size adequate to support the measurement of outcomes?
   - If different interventions, exposures, factors or events are being compared, the number of subjects studied must be adequate for the effects of the independent variable (what is changed) on the dependent variable (what is measured). This "number" depends on the magnitude of the difference in outcome between interventions or exposure, the probability that differences could be attributed to chance (p value) or random variation and the characteristics of the data, including the frequency of the event (outcome) in the described population. Typically, this is a difficult number for the non-expert to estimate, and the "answer" or rationale should be noted by the investigator experts (the authors). Often, the methods by which the outcomes are measured or described (as well as validated and referenced in the paper) give a good indication of a minimum number of subjects "needed" or expected. For instance, a diagnostic test will reference the sensitivity of the test and the investigator using the diagnostic test should "measure" at least the minimum number of subjects (phew!). Do the investigators provide this type of information or references?

e) Sample / Population size: Is the sample size (population numbered) large enough to analyze statistically?
   - Case reports and case studies describe one patient. This is not a large enough sample size to analyze statistically. Therefore the appropriate "answer" for a case study would be that case studies do not use statistical analyses. Subject descriptions should be richly detailed.
   - A general rule of thumb for clinical trials is a study size of 30 patients with at least one point of interaction or intervention. Therefore, in a clinical trial comparing 2 interventions, one would look for at least 30 study participants per group. A case series typically should have at least 10 subjects with well defined characteristics, baseline differences and medical histories which affect the outcome. Fewer participants might be used, and authors should describe their rationale for using a particular sample size – see 3f. (Again, non-experts should be able to depend on the investigator expert for information, evidence and rationale.)

f) Sample / Population size: Was there rationale for choosing this sample size?

g) Follow-up / Accountability
   i) Were all study participants or subjects accounted for at the end of the study?
   ii) Are the reasons why patients withdraw from clinical trials included in the follow-up information?
   iii) Were all participants accounted for at the end of the study? Was any drop-out discussed? A rule of thumb is that a drop-out of more than 20% generally threatens validity. Did authors note and analyze all patient results with the "intention to treat," that is, including all patients in the group to which they were assigned regardless of whether or not they finished the study?
   iv) How much the loss of patients to the study end or follow-up threatens validity depends on the magnitude of the treatment effect. Authors should determine this by looking at worse case scenarios: all patients lost to follow-up treatment groups had a negative outcome, and all patents in the control arm had positive outcome. Authors should calculate results using all patients theoretically completing the study in the group to which they were assigned. (Note for section 9a "Validity.")

h) Ethical Approval
   i) If appropriate, was ethical approval of the participants and institutional review board obtained? This is different from association and financial disclosures (see section 12).

4) Methods: Intervention

a) Description of Intervention
   i) What specific intervention or other maneuver was given?
   ii) Is the intervention sufficiently described so that the reader (practitioner) could adequately deliver the same intervention?

b) Comparison / Placebo / Control Intervention
   i) With what was the investigated or experimental intervention compared?
   ii) Was a placebo, "sham" or a "zero" utilized? Was it sufficiently similar to the real treatment that the study participant was unaware (or blind) to whether the treatment was real or not?
      - A placebo is an inactive substance or intervention given to a subject in place of an experimental substance (intervention) used as a control in an experiment to test the effectiveness of the experimental substance. In a clinical study, placebo should mimic the experimental substance so exactly (taste, odor, shape, feel, look, administration, etc.) that the subjects and the administrator should not be able to tell the difference between the placebo and the experimental substance.
      - A "sham" usually refers to a maneuver that takes the place of the experimental intervention. Shams are used in place of acupuncture, chiropractic, and massage techniques as well as psychological assessments, surveys, etc. Shams are used as controls or placebos.
      - A "zero" point is a baseline point or assessment where initial subject data is taken, but no experimental, control, placebo or sham intervention is given.
   iii) Are any of the comparison interventions considered "standards?" Have they been "validated" in other studies? Are references given?
      - A standard is a well accepted, usually validated outcome.
      - "Validated" outcomes, assessments, measurement tools or tests have solid experimental evidence that substantiates their use and acceptance. A clinical research paper should provide information and evidence with references for outcomes and measurements in the introduction or overview section. References should be provided.

c) Randomization
   - Randomization: subjects are assigned to a specific group (intervention) by a disciplined process that ensures that each subject has an equal chance of being assigned to a group. Subjects in one group are as likely to possess a characteristic as those in another group.
   - Random selection: Subjects are randomly chosen (selected) from a population to be in a research study (e.g., to receive a survey).
   - Random allocation: Subjects included or enrolled in a study (sample of a population) are randomly placed (allocated) into one study group (intervention) or another. The sample may have defined inclusion / exclusion criteria in order to be included or enroll in the study.
   - Selection bias occurs when differences in the comparison groups are attributable to incomplete randomization. Subjects may differ in additional factors or aspects not under study that may affect the outcome of the disease (confounders). For example, if a study indicates coffee drinkers have a higher rate of coronary artery disease than non-coffee drinkers and further analysis shows that the non-coffee drinkers exercised substantially more than the coffee drinkers, coffee drinking would be a confounding factor since other substantial studies show exercise has a direct association with coronary artery disease. Selection bias is an issue when patients are chosen for an observation (or intervention), and must be dealt with in the design of a study (randomization or intentional matching on confounding factors). If these confounding factors can be controlled, they are referred to as "controlled" or "constant variables." In the example, the amount of exercise could be controlled at a specific level while the amount of coffee consumption is specifically varied.

   i) With this study design, are various groups, controls and randomization possible?
   ii) Was there a control population or a comparison group identical in all aspects other than the intervention or exposure?
      - Control groups should be identified in RCTs. Case-control and cohort studies may include true "zero" control groups (no intervention or exposure) or a comparison group. Case series, case studies and case reports do not include a true control or comparison. Authors using these study designs may analyze patient histories before interventions or exposures, designating a baseline or a "zero point." Surveys usually do not include a control, although after initial analysis, the investigator may designate a group as a "control" based on a specific demographic or an exposure / non-exposure determined from the survey, etc.

iii) Was there randomization as far as who received the intervention? How was this assured?
- Centralized computer randomization using randomly generated numbers which are then concealed (as in a sealed envelope) from both the study participant and the person administering recruitment and eligibility (as well as to those who measure the outcome and assess the results) is ideal and is often used in multi-centered trials. Smaller trials may use an independent person (e.g., the hospital pharmacy) to "allocate" the randomization.
- Not all clinical trials include randomization.

iv) If there were no groups (case study, case series, community or cross-sectional studies) or randomization, how were participants selected and did the author protect against selecting participants that would respond well to the treatment?

d) Blinding: study subjects / participants (masking)
i) Was the study population / subjects / participants blind to the type of intervention (experimental, sham or placebo) or exposure?
ii) Could participants tell if they received the full intervention or were in a control group?
iii) Were experimental and control groups (if used) treated equally?

e) Blinding: investigators, researchers, administrators selecting, randomizing or allocating participants, dispensing intervention(s), assessing outcomes and analyzing results
i) Were those involved in determining the effect of the intervention or exposure on the outcomes and analysis or results blinded?
ii) Were the "interveners" or delivery practitioners "blind" to what the patients were being given?
iii) Were those analyzing the results blind to the population, initial / baseline measurements, the intervention (or placebo) given, and the on-going results of the experiment?
iv) Were there adequate study personnel and were responsibilities well defined?

5) Methods: Outcomes, measurement, observation

a) Outcomes measured:
- "Outcome" refers to the clinical event of interest, desired effect, end product or consequence following an intervention or exposure that is measured in some way. "Outcomes" and results are different: the measurement of an outcome is reported as a result.
- Outcomes should be clinically relevant such as "a reduction in blood pressure," "reduced mortality," "better quality of life," "management of blood glucose levels," "resolution of pain," etc.
- Biologic outcomes or surrogate endpoints (decrease in blood glucose levels, decrease in serum IgE levels, half-life of a drug in serum samples) do not necessarily singularly correlate with a clinical outcome (control of diabetes, death, recovery from a disease, decrease in blood pressure). Sometimes, several biological outcomes are assessed in order to measure a clinical outcome quantitatively. A typical flaw in clinical research is to make a "claim" regarding a clinical outcome when a biological outcome or surrogate endpoint is assessed, especially when there are several biological outcomes or surrogates should serve as markers for the clinical status of the subject.
- Changes in measurements of outcomes should reflect a change in the status of a patient.

i) What outcome(s) were measured? Describe or list the outcomes that were measured or the investigators desired to accomplish. Were the outcomes desired (purpose) the same as what was measured? (Are the definitions of the outcomes clear and logical?)
ii) Are the outcomes measured or observed clinically relevant?
- If biologic outcomes or surrogate endpoints (above) are used to quantify an outcome, has the biologic outcome or surrogate endpoint been shown to be directly applicable to the clinical outcome? Are references given?

iii) How were the outcomes measured (or observed)?
iv) For cohort studies, were outcomes defined at the start of the study?
v) For case-control studies, did the exposure precede the outcome?
vi) Was the study of adequate length for experimentation / observation / trial and measurement of the outcome(s)?
- Were an adequate number of visits provided at appropriate intervals and frequency?

b) Are outcomes valid?
   - Validity in terms of "outcomes"
   i) Are these outcomes standard or typical outcomes desired for the patient health issue, typical of a comparison intervention or typical of a comparative exposure?
   ii) Were the outcomes objectively set (e.g., compared to or referenced in other studies)?

c) Does the measurement "tool" match the outcome (for example, a self-directed survey versus a measurement of function, metabolic marker, etc.)? Are there alternative measurements available or used in this study?
   i) Are there alternative measurements which are considered "gold standards," or standards?
   ii) Is the endpoint measurement clinically relevant?
   iii) Is the measurement or measurement tool validated, well known and accepted (are there references indicated)?
   iv) Is there an explanation of why that outcome was selected as well as the measurement?

d) Are measurements reliable?
   - Reliability is the extent to which a repeated measurement of a defined, stable event or phenomenon has similar results under stable circumstances (i.e., repeatability). It may be reported as "precision."

e) Are measurements reproducible?
   - Reproducibility is that aspect of reliability that demonstrates that a repeated measurement of a stable event or phenomenon will have the same result when measured by different people, instruments, test or tool. Reproducibility adds strength to the validity of a measurement. Consider whether the measurement could be conducted by a different practitioner in a different setting, or if it is so specialized that it needs special instruments, people or environment. Reproducibility implies that the measurement is reliable (repeatable, precise) and accurate, although if a reproducible error also occurs with each measurement, the measurement may be reproducible, but the consistent error may make it inaccurate. (Accuracy refers to how close a repeated measurement is to the true, theoretical value, or how "on-target" a measurement is.)

6) Methods: Analysis and Validity

a) Is the analysis valid and sufficient?
   i) Are the analytical or statistical methods used to analyze the outcome measure(s) appropriate, recognized or well-known, sufficiently described and sufficiently explained?
   ii) Are any statistical methods or tools validated by other studies? Are references given?

b) Is this trial or experiment also being used to test new methods of measurement and analysis? (If so, are sufficient methods in place to accurately assess the new measure?)

7) Key results

a) Are the basic data adequately described?
   i) Are results presented clearly, objectively and in sufficient detail to enable the reader to draw their own conclusions?

b) Describe key results (as bullet points and adequate key word phrases).

c) Are adverse effects adequately described and explained?

d) Are tables and graphs readable and clear?

e) Does the data presented match the text (descriptions, conclusions, discussion)?

8) Statistical significance

a) Describe the statistical methods. Do they seem adequate for this study design?
   i) Are the methods used to analyze the data appropriate, recognized or well-known, sufficiently described, explained or referenced?

b) Have "p values" been calculated?
   - "p value" is the probability that any particular outcome would have arisen by chance.
   - "Standard scientific practice" (often considered somewhat arbitrary):
     - $p < 0.05$ (p value less than one in twenty) is "statistically significant"
     - $p < 0.01$ (p value less than one in one hundred) is "statistically highly significant"
     - the study should state the cutoff chosen for the study (e.g., $p < 0.05$ or $p < 0.01$)
     - a statistically significant result suggests that the authors should reject the "null hypothesis;" that is, the hypothesis that there is no real difference between study groups
     - p values greater than 0.05 (e.g., 0.49, or 0.30) are not considered statistically significant
     - p values in the non-significant range indicate that either there is not difference between groups OR there are too few subjects to demonstrate a difference (if a difference exists). It does not determine which circumstance the p value reflects.
     - Typically, "positive trials" show a statistically significant difference between groups or arms of a trial, and "negative trials" appear to show no significant difference between groups or arms.

c) Have "p values" been stated and interpreted appropriately?
   i) What level of difference between the groups, outcomes or interventions constitutes a statistically significant effect according to the study designers / authors?
   ii) Does the calculated value match the conclusion of the authors?
   iii) Does the statistical significance match clinical significance?

d) Have confidence intervals (CI) been calculated, stated correctly, matching the data, and have they been interpreted adequately and correctly? If so, what are they?
   - The confidence interval quantifies the uncertainty in a measurement, It defines the "% confidence" that the true value of a measurement or calculation lies within a certain range, allows the estimation for both positive trials (show a statistically significant difference between groups or arms of a trial) and negative trials (those which appear to show no significant difference) whether the strength of the evidence (results of outcomes measured) is strong or weak, and whether the study is definitive (precludes the need for further, similar or repeated studies).
   - The confidence interval states an upper and lower limit and the likelihood that a certain percentage of the results will fall between that interval. For example, a typical clinically relevant confidence interval of 95%.
   - A "95% confidence interval" means that there is a 95% chance that the real difference between 2 groups would fall between the upper and lower limits measured.
   - The wider the confidence interval, the more likely that a certain result will fall within that interval. Strong evidence will have a wider confidence interval.
   - Look for statements like, "In a trial for treatment of back pain, 42.6% of patients randomized to Intervention A returned to work between 2 and 4 weeks. For patients randomized to treatment B, 32.4% of patients returned to work in 2 to 4 weeks. The point estimate of the difference between groups (or treatment effect) is 10.2%; 95% CI (confidence interval) 5.9-14.5.

e) Do the authors' conclusions and discussion reflect the calculated statistics, including the mean, p values or confidence intervals (if calculated and reported)?

9) Overall, is the study valid?

- Validity is the strength of the conclusions, inferences or propositions providing the best estimation of the truth or falsity of a hypothesis, proposition or conclusion. Validity is the absence of bias – and forms of bias.
- Internal Validity: Baseline characteristics of a group are the same, except for the characteristic, exposure, intervention under study and study groups are comparable. An internally valid study demonstrates a relationship between the intervention and the outcome measured and determines if it a causal relationship? For example, did the intervention cause a decrease in hypertension? Did the exposure cause a disease?
- External validity: Baseline characteristics of the entire population (population sample) under study are similar to general population. Since the study group and general population are comparable, the results can be generalized to the entire population.
- Was the study design (construction) adequate to allow measurement of the outcome without other issues confounding results or causing some aspect of bias?
- Can the results of the study be applied to other settings (e.g., a study carried out in a large hospital setting repeated on a single patient in an urban, private practice) or patient populations?

a) Assess the over all validity of the study considering the aspects already appraised in this CAT. *Summarize* any threats to validity or flaws in bullets or phrases using the guide questions below and your appraisal of the key elements of this paper. Note particularly noteworthy (to you) solutions the authors included in their study design to address potential threats to validity.

i) Was a placebo, sham or zero treatment used? Was it adequate?
ii) Was there a comparison intervention or sample group? Was it well defined and adequate?
iii) Was the randomization or selection "blinded" or concealed from:
   - The participants being studied. Were study participants randomized into groups or were they allocated depending on a particular characteristic (e.g., were they exposed to environmental pollutants or not and thus allocated to "exposed" and "not exposed" groups?
     - If patients cannot be "blinded" to a placebo, sham or control intervention, the determination of the outcome(s) should not depend on subjective patient measurements or responses.
   - The person determining eligibility (so that not only people who would respond to the treatment were selected for inclusion)?
   - Those assessing / measuring the outcomes?
   - Those analyzing the outcome measurements and determining the results?
   - Was the provider delivering the intervention different from the person assessing the outcome and analyzing results?

iv) Follow-up / intention to treat: All people allocated to each arm of the treatment regimen or intervention should analyzed together as representing that treatment, whether or not they received or completed the treatment. Failure to complete the analysis with "intention to treat" defeats the purpose of random allocation. Were patients analyzed in the groups to which they were randomized or allocated?
v) Was follow-up adequate? Complete? (Assessed in 3g)

b) Was systematic bias avoided or minimized?

- Systematic bias is anything which erroneously influences the conclusions about groups and distorts comparisons. Systematic bias is assessed in relation to RCTs, some (non-randomized, controlled) clinical trials, cohort studies and case control studies. Groups should as similar as possible except for the factor being considered or studied (reception of an intervention, exposure or an agent or event, etc.). Groups should receive the same (equal not equitable) intervention, exposure, contact and information, assessment using the same outcome measures and should be analyzed using the same methods and tools.
  - RCTs: systematic bias can theoretically be avoided by selecting a sample of participants from a particular population and allocating them randomly into different groups. Sources of systematic bias in RCTs include:
    - Selection bias refers to differences in the comparison groups attributable to incomplete randomization.
    - Performance bias refers to systematic differences in the care provided to groups other than the intervention or exposure.

- Participation bias occurs when some members of a group refuse to participate or do not finish the study because of some difference (magnitude of the illness, exposure, reaction to or attitude toward the intervention, ability to participate such as disability or age, etc.)
- Exclusion bias refers to systematic differences in withdrawal / drop-out from the study.
- Detection bias refers to differences in the outcome assessment.

- Clinical trials: one major source of systematic bias is the comparison of two groups which have inherent, self-selected differences before the intervention or exposure.
- Cohort studies:
  - Selection bias: do the groups differ in some important characteristic other than th exposure being studied or measured? Was the identification and selection of a comparable control group logical and valid (equal in age, gender, culture, socioeconomic status, etc.) so that the single difference between groups is the exposure to the agent studied?
  -
- Case-control studies: the experience of individuals with and without a particular disease analyzed retrospectively to identify putative (expected, presupposed) causative agents, with a particular, singular identification of when a patient becomes a "case."

c) Were the outcomes measured the same as those stated in the study objectives? Were the outcomes appropriately described in the title of the paper?

10) Clinical significance

a) Are the outcomes and results what the patient and the health care provider are interested in?

b) Does the statistical significance match clinical significance?

11) Author's conclusions

a) Have the authors expressed the effects of an intervention or exposure in terms of the likely benefit or harm which an individual patient can expect?

b) Do the authors' conclusions match the data, analysis and statistical significance?

12) Statements of affiliation, association, financial support or influence / Disclaimers

- Authors should reveal their associations, whether this research has been supported financially (financial disclosures) and whether the authors or associates have been compensated by an institution, company or individual who stands to gain something by the success or failure of the study or by the dissemination of the evidence. Examples of claim of affiliation, association and financial support might read:

1. Obtained funding: Weinstein, T. Tosteson, A. Tosteson. Administrative, technical, or material support: Weinstein, Lurie, T. Tosteson. Study supervision: Weinstein, Lurie.

2. Financial Disclosures: Dr Weinstein reports that he is Editor-in-Chief of Spine, has been a consultant to United Health Care (proceeds are donated to the Brie Fund, a fund for children with disabilities in the name of his daughter who passed away from leukemia), and has been a consultant to the Foundation for Informed Medical Decision-Making, proceeds to the Department of Orthopaedics, Dartmouth. Dr Lurie reports that he receives grant support from St Francis Medical Technologies and American Board of Orthopaedic Surgery; has served on advisory boards for Ortho-MacNeil Pharmaceuticals, the Robert Graham Center of the American Academy of Family Practice, Pfizer, and Centocor; and as a consultant for Myexpertdoctor.com, Pacific Business Group on Health, and the Foundation for Informed Medical Decision- Making. Mr Hanscom reports working for the National Spine Network and receiving funding from Medtronic.

   -- Weinstein JAMA, November 22/29, 2006—Vol 296, No. 20.

13) Comments and notes:

a) Strengths of this study (include in a "Best Evidence for a Topic" (BET) summary)

b) Limitations of this study (include in a "Best Evidence for a Topic" (BET) summary)

c) Does this new research add to the literature in any way?

d) After the Critical Appraisal, I rate the evidence in this paper (very strong, strong, moderate, weak, very weak):

14) Clinical Impact statement

- Two to three sentences summarizing your perception of the potential impact on your health care profession (chiropractic medicine, acupuncture, oriental medicine, naturopathic medicine, massage therapy, etc.), especially on *your* practice.
- Comment on: "Does this paper help me answer my clinical, patient oriented question or problem?"

Resources and References

BestBETS.org Emergency Department at Manchester Royal Infirmary UK BestBETS Best Evidence Tables About Participation Appraising the Literature http://www.bestbets.org/home/participation.php#4 (accessed 17 July 2008).

BestBETS.org Emergency Department at Manchester Royal Infirmary UK BestBETS Best Evidence Tables Resources CA Worksheets http://www.bestbets.org/links/BET-CA-worksheets.php (accessed 17 July 2008).

Center for Evidence Based Medicine, EBM Tools, Critical Appraisal CAT Maker http://www.cebm.net/index.aspx?o=1216 (accessed 17 July 2008).

Centre for Health Evidence (University of Alberta CA) Users' Guides for Evidence-Based Practice http://www.cche.net/usersguides/main.asp (accessed 17 July 2008).

Cote L, Turgeon J. Appraising qualitative research articles in medicine and medical education. Medical Teacher 2005;27(1):71-5.

GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ 2004 Jun 19;328(7454):1490-7.

Greenhalgh T. How to read a paper: the basics of evidence-based medicine (3rd ed). London: Wiley, 2006.

ISBN 1405139765.

Guyatt G, Rennie D. User's guide to the medical literature: A manual for evidence-based clinical practice. Chicago: AMA Press, 2002. ISBN 1579471919.

Haneline M. Evidence-based chiropractic practice. Boston Toronto London Singapore: Jones & Bartlett Publishers, 2007. ISBN 076373571X

Haneline M., Cooperstein R. Appraisal of journal articles: asking the right questions. JACA 2006 May/Jun.

Sackett, DL. Richardson WS, Rosenberg W, Haynes RB. Evidence-based Medicine: How to Practice and Teach EBM. 2nd edition. Churchill Livingtone, 2000.

Strauss SE, Richardson WS, Glasziou P, Haynes RB. Evidence-based medicine (3rd ed.). Edinburgh New York: Elsevier/Churchill Livingstone, 2005.