

## **CHECKLIST FOR EVALUATING A RESEARCH REPORT**

**Provided by Dr. Blevins**

### **1. The Title**

- a. Is it clear and concise?
- b. Does it promise no more than the study can provide?

### **INTRODUCTION**

#### **2. The Problem**

- a. It is clearly stated?
- b. Is it properly defined?
- c. Is its significance recognized?
- d. Are specific questions raised; hypotheses clearly stated?
- e. Are assumptions and limitations stated?
- f. Are important terms defined?

#### **3. Review of Related Literature**

- a. Is it adequately covered?
- b. Are important findings noted?
- c. Is it well organized?
- d. Is an effective summary provided?

### **METHODOLOGY**

#### **4. Procedures Used**

- a. Is the research design described in detail?
- b. Is it adequate?
- c. Are the samples described?
- d. Are relevant variables recognized?
- e. Are appropriate controls provided?
- f. Are data-gathering instruments appropriate?
- g. Are validity and reliability established?
- h. Is the statistical treatment appropriate?

### **RESULTS**

#### **5. Data Analysis**

- a. Is appropriate use made of tables and figures?
- b. Is the textual discussion clear and concise?
- c. Is the analysis of data relationships logical and perceptive?
- d. Is the statistical analysis accurately interpreted?

### **DISCUSSION/CONCLUSION**

#### **6. Summary and Conclusions**

- a. Is the problem restated?
- b. Are the procedures and findings concisely presented?
- c. Is the analysis objective?
- d. Are the findings and conclusions justified by the data presented and analyzed?

## **STEPS IN ANALYZING A RESEARCH ARTICLE**

**Provided by Dr. Blevins**

### **INTRODUCTION**

- Does it properly introduce the subject?
- Does it clearly state the purpose of what is to follow?
- Does it briefly state why this report is different from previous publications?

### **METHODS AND MATERIALS**

- Is the test population clearly stated? Is it appropriate for the experiment? Should it be larger? More comprehensive?
- Is the control population clearly stated? Are all variables controlled? Should it be larger? More comprehensive?
- Are methods clearly described or referenced so the experiment could be repeated?
- Are materials clearly described and when appropriate, manufacturers footnoted?
- Are all statements and descriptions concerning design of test and control populations and materials and methods included in this section?

### **RESULTS**

- Are results for all parts of the experimental design provided?
- Are they clearly presented with supporting statistical analyses and/or charts and graphs when appropriate?
- Are results straightforwardly presented without a discussion of why they occurred?
- Are all statistical analyses appropriate for the situation and accurately performed?

### **DISCUSSION**

- Are all results discussed?
- Are all conclusions based on sufficient data?
- Are appropriate previous studies integrated into the discussion section?

### **ABSTRACT**

- Does the first sentence contain a clear statement of the purpose of the article (without starting....The purpose of this article is to.....)
- Is the test population briefly described?
- Does it conclude with a statement of the experiment's conclusions?

**STANDARD QUALITY ASSESSMENT CRITERIA FOR EVALUATING  
PRIMARY RESEARCH PAPERS FROM A VARIETY OF FIELDS**

**Prepared by:**

**Leanne M. Kmet, Robert C. Lee and Linda S. Cook**

**Quantitative Studies Questions**

1. Question / objective sufficiently described?
2. Study design evident and appropriate?
3. Method of subject/comparison group selection or source of information/input variables described and appropriate?
4. Subject (and comparison group, if applicable) characteristics sufficiently described?
5. If interventional and random allocation was possible, was it described?
6. If interventional and blinding of investigators was possible, was it reported?
7. If interventional and blinding of subjects was possible, was it reported?
8. Outcome and (if applicable) exposure measure(s) well defined and robust to measurement / misclassification bias? means of assessment reported?
9. Sample size appropriate?
10. Analytic methods described/justified and appropriate?
11. Some estimate of variance is reported for the main results?
12. Controlled for confounding?
13. Results reported in sufficient detail?
14. Conclusions supported by the results?

**Manual for Quality Scoring of Quantitative Studies  
Definitions and Instructions for Quality Assessment Scoring**

How to calculate the summary score:

Total sum = (number of "yes" \* 2) + (number of "partials" \* 1)

Total possible sum = 28 – (number of "N/A" \* 2)

Summary score: total sum / total possible sum

**1. Question or objective sufficiently described?**

- **Yes:** Is easily identified in the introductory section (or first paragraph of methods section). Specifies (where applicable, depending on study design) all of the following: purpose, subjects/target population, and the specific intervention(s) /association(s)/descriptive parameter(s) under investigation. A study purpose that only becomes apparent after studying other parts of the paper is not considered sufficiently described.
- **Partial:** Vaguely/incompletely reported (e.g. “describe the effect of” or “examine the role of” or “assess opinion on many issues” or “explore the general attitudes” ...); or some information has to be gathered from parts of the paper other than the introduction/background/objective section.
- **No:** Question or objective is not reported, or is incomprehensible.
- **N/A:** Should not be checked for this question.

## 2. Design evident and appropriate to answer study question?

- **Note:** If the study question is not given, infer from the conclusions
- **Yes:** Design is easily identified and is appropriate to address the study question / objective.
- **Partial:** Design and /or study question not clearly identified, but gross inappropriateness is not evident; or design is easily identified but only partially addresses the study question.
- **No:** Design used does not answer study question (e.g., a comparison group is required to answer the study question, but none was used); or design cannot be identified.
- **N/A:** Should not be checked for this question.

## 3. Method of subject selection (and comparison group selection, if applicable) or source of information/input variables (e.g., for decision analysis) is described and appropriate.

- **Yes:** Described and appropriate. Selection strategy designed (i.e., consider sampling frame and strategy) to obtain an unbiased sample of the relevant target population or the entire target population of interest (e.g., consecutive patients for clinical trials, population-based random sample for case-control studies or surveys). Where applicable, inclusion/exclusion criteria are described and defined (e.g., “cancer” -- ICD code or equivalent should be provided). Studies of volunteers: methods and setting of recruitment reported. Surveys: sampling frame/ strategy clearly described and appropriate.
- **Partial:** Selection methods (and inclusion/exclusion criteria, where applicable) are not completely described, but no obvious inappropriateness. Or selection strategy is not ideal (i.e., likely introduced bias) but did not likely seriously distort the results (e.g., telephone survey sampled from listed phone numbers only; hospital based case-control study identified all cases admitted during the study period, but recruited controls admitted during the day/evening only). Any study describing participants only as “volunteers” or “healthy volunteers”. Surveys: target population mentioned but sampling strategy unclear.

- **No:** No information provided. Or obviously inappropriate selection procedures (e.g., inappropriate comparison group if intervention in women is compared to intervention in men). Or presence of selection bias which likely seriously distorted the results (e.g., obvious selection on “exposure” in a case-control study).
- **N/A:** Descriptive case series/reports.

**4. Subject (and comparison group, if applicable) characteristics or input variables/information (e.g., for decision analyses) sufficiently described?**

- **Yes:** Sufficient relevant baseline/demographic information clearly characterizing the participants is provided (or reference to previously published baseline data is provided). Where applicable, reproducible criteria used to describe/categorize the participants are clearly defined (e.g., ever-smokers, depression scores, systolic blood pressure > 140). If “healthy volunteers” are used, age and sex must be reported (at minimum). Decision analyses: baseline estimates for input variables are clearly specified.
- **Partial:** Poorly defined criteria (e.g. “hypertension”, “healthy volunteers”, “smoking”). Or incomplete relevant baseline / demographic information (e.g., information on likely confounders not reported). Decision analyses: incomplete reporting of baseline estimates for input variables.
- **No:** No baseline / demographic information provided. Decision analyses: baseline estimates of input variables not given.
- **N/A:** Should not be checked for this question.

**5. If random allocation to treatment group was possible, is it described?**

- **Yes:** True randomization done - requires a description of the method used (e.g., use of random numbers).
- **Partial:** Randomization mentioned, but method is not (i.e. it may have been possible that randomization was not true).
- **No:** Random allocation not mentioned although it would have been feasible and appropriate (and was possibly done).
- **N/A:** Observational analytic studies, uncontrolled experimental studies, surveys, descriptive case series / reports, decision analyses

**6. If interventional and blinding of investigators to intervention was possible, is it reported?**

- **Yes:** Blinding reported.
- **Partial:** Blinding reported, but it is not clear who was blinded.
- **No:** Blinding would have been possible (and was possibly done) but is not reported.
- **N/A:** Observational analytic studies, uncontrolled experimental studies, surveys, descriptive case series and reports, decision analyses.

## 7. If interventional and blinding of subjects to intervention was possible, is it reported?

- **Yes:** Blinding reported.
- **Partial:** Blinding reported but it is not clear who was blinded.
- **No:** Blinding would have been possible (and was possibly done) but is not reported.
- **N/A:** Observational studies, uncontrolled experimental studies, surveys, descriptive case series / reports.

## 8. Outcome and (if applicable) exposure measure(s) well defined and robust to measurement / misclassification bias? Means of assessment reported?

- **Yes:** Defined (or reference to complete definitions is provided) and measured according to reproducible, “objective” criteria (e.g., death, test completion – yes/no, clinical scores). Little or minimal potential for measurement / misclassification errors. Surveys: clear description (or reference to clear description) of questionnaire/interview content and response options. Decision analyses: sources of uncertainty are defined for all input variables.
- **Partial:** Definition of measures leaves room for subjectivity, or not sure (i.e., not reported in detail, but probably acceptable). Or precise definition(s) are missing, but no evidence or problems in the paper that would lead one to assume major problems. Or instrument/mode of assessment(s) not reported. Or misclassification errors may have occurred, but they did not likely seriously distort the results (e.g., slight difficulty with recall of long-ago events; exposure is measured only at baseline in a long cohort study). Surveys: description of questionnaire/interview content incomplete; response options unclear. Decision analyses: sources of uncertainty are defined only for some input variables.
- **No:** Measures not defined, or are inconsistent throughout the paper. Or measures employ only ill-defined, subjective assessments, e.g. “anxiety” or “pain.” Or obvious misclassification errors / measurement bias likely seriously distorted the results (e.g., a prospective cohort relies on self-reported outcomes among the “unexposed” but requires clinical assessment of the “exposed”). Surveys:
  - No description of questionnaire/interview content or response options. Decision analyses: sources of uncertainty are not defined for input variables.
  - **N/A:** Descriptive case series / reports.

## 9. Sample size appropriate?

- **Yes:** Seems reasonable with respect to the outcome under study and the study design. When statistically significant results are achieved for major outcomes, appropriate sample size can usually be assumed, unless large standard errors ( $SE > 1/2$  effect size) and/or problems with multiple testing are evident. Decision analyses: size of modeled cohort / number of iterations specified and justified.

- **Partial:** Insufficient data to assess sample size (e.g., sample seems “small” and there is no mention of power/sample size/effect size of interest and/or variance estimates aren’t provided). Or some statistically significant results with standard errors  $> 1/2$  effect size (i.e., imprecise results). Or some statistically significant results in the absence of variance estimates. Decision analyses: incomplete description or justification of size of modeled cohort / number of iterations.
- **No:** Obviously inadequate (e.g., statistically non-significant results and standard errors  $> 1/2$  effect size; or standard deviations  $> \_$  of effect size; or statistically non-significant results with no variance estimates and obviously inadequate sample size). Decision analyses: size of modeled cohort / number of iterations not specified.
- **N/A:** Most surveys (except surveys comparing responses between groups or change over time). Descriptive case series / reports.

#### 10. Analysis described and appropriate?

- **Yes:** Analytic methods are described (e.g. “chi square”/ “t-tests”/“Kaplan-Meier with log rank tests”, etc.) and appropriate.
- **Partial:** Analytic methods are not reported and have to be guessed at, but are probably appropriate. Or minor flaws or some tests appropriate, some not (e.g., parametric tests used, but unsure whether appropriate; control group exists but is not used for statistical analysis). Or multiple testing problems not addressed.
- **No:** Analysis methods not described and cannot be determined. Or obviously inappropriate analysis methods (e.g., chi-square tests for continuous data, SE given where normality is highly unlikely, etc.). Or a study with a descriptive goal / objective is over-analyzed.
- **N/A:** Descriptive case series / reports.

#### 11. Some estimate of variance (e.g., confidence intervals, standard errors) is reported for the main results/outcomes (i.e., those directly addressing the study question/ objective upon which the conclusions are based)?

- **Yes:** Appropriate variances estimate(s) is/are provided (e.g., range, distribution, confidence intervals, etc.). Decision analyses: sensitivity analysis includes all variables in the model.
- **Partial:** Undefined “+/-“ expressions. Or no specific data given, but insufficient power acknowledged as a problem. Or variance estimates not provided for all main results/outcomes. Or inappropriate variance estimates (e.g., a study examining change over time provides a variance around the parameter of interest at “time 1” or “time 2”, but does not provide an estimate of the variance around the difference). Decision analyses: sensitivity analysis is limited, including only some variables in the model.
- **No:** No information regarding uncertainty of the estimates. Decision analyses: No sensitivity analysis

- **N/A:** Descriptive case series / reports. Descriptive surveys collecting information using open-ended questions.

## 12. Controlled for Confounding?

- **Yes:** Randomized study, with comparability of baseline characteristics reported (or non-comparability controlled for in the analysis). Or appropriate control at the design or analysis stage (e.g., matching, subgroup analysis, multivariate models, etc). Decision analyses: dependencies between variables fully accounted for (e.g., joint variables are considered).
- **Partial:** Incomplete control of confounding. Or control of confounding reportedly done but not completely described. Or randomized study without report of comparability of baseline characteristics. Or confounding not considered, but not likely to have seriously distorted the results. Decision analyses: incomplete consideration of dependencies between variables.
- **No:** Confounding not considered, and may have seriously distorted the results. Decision analyses: dependencies between variables not considered.
- **N/A:** Cross-sectional surveys of a single group (i.e., surveys examining change over time or surveys comparing different groups should address the potential for confounding). Descriptive studies: Studies explicitly stating the analysis is strictly descriptive/exploratory in nature.

## 13. Results reported in sufficient detail?

- **Yes:** Results include major outcomes and all mentioned secondary outcomes.
- **Partial:** Quantitative results reported only for some outcomes. Or difficult to assess as study question/objective not fully described (and is not made clear in the methods section), but results seem appropriate.
- **No:** Quantitative results are reported for a subsample only, or “n” changes continually across the denominator (e.g., reported proportions do not account for the entire study sample, but are reported only for those with complete data i.e., the category of “unknown” is not used where needed). Or results for some major or mentioned secondary outcomes are only qualitatively reported when quantitative reporting would have been possible (e.g., results include vague comments such as “more likely” without quantitative report of actual numbers).
- **N/A:** Should not be checked for this question.

## 14. Do the results support the conclusions?

- **Yes:** All the conclusions are supported by the data (even if analysis was inappropriate). Conclusions are based on all results relevant to the study question, negative as well as positive ones (e.g., they aren’t based on the sole significant finding while ignoring the negative results). Part of the conclusions may expand beyond the results, if made in addition to rather than instead of those strictly supported by data, and if including indicators of their interpretative nature (e.g., “suggesting,” “possibly”).

- **Partial:** Some of the major conclusions are supported by the data, some are not. Or speculative interpretations are not indicated as such. Or low (or unreported) response rates call into question the validity of generalizing the results to the target population of interest (i.e., the population defined by the sampling frame/strategy).
- **No:** None or a very small minority of the major conclusions are supported by the data. Or negative findings clearly due to low power are reported as definitive evidence against the alternate hypothesis. Or conclusions are missing. Or extremely low response rates invalidate generalizing the results to the target population of interest (i.e., the population defined by the sampling frame/ strategy).
- **N/A:** Should not be checked for this question.

## Checklist for assessing the quality of qualitative studies

1. Question / objective sufficiently described?
2. Study design evident and appropriate?
3. Context for the study clear?
4. Connection to a theoretical framework / wider body of knowledge?
5. Sampling strategy described, relevant and justified?
6. Data collection methods clearly described and systematic?
7. Data analysis clearly described and systematic?
8. Use of verification procedure(s) to establish credibility?
9. Conclusions supported by the results?
10. Reflexivity of the account?

## Manual for Quality Scoring of Qualitative Studies Definitions and Instructions for Quality Assessment Scoring

### How to calculate the summary score

Total sum = (number of “yes” \* 2) + (number of “partials” \* 1)

Total possible sum = 20

Summary score: total sum / total possible sum

### Quality assessment

#### 1. Question / objective clearly described?

- **Yes:** Research question or objective is clear by the end of the research process (if not at the outset).
- **Partial:** Research question or objective is vaguely/incompletely reported.
- **No:** Question or objective is not reported, or is incomprehensible.

#### 2. Design evident and appropriate to answer study question?

Note: If the study question is not clearly identified, infer appropriateness from results/conclusions.

- **Yes:** Design is easily identified and is appropriate to address the study question.
- **Partial:** Design is not clearly identified, but gross inappropriateness is not evident; or design is easily identified but a different method would have been more appropriate.
- **No:** Design used is not appropriate to the study question (e.g. a causal hypothesis is tested using qualitative methods); or design cannot be identified.

### 3. Context for the study is clear?

- **Yes:** The context/setting is adequately described, permitting the reader to relate the findings to other settings.
- **Partial:** The context/setting is partially described.
- **No:** The context/setting is not described.

### 4. Connection to a theoretical framework / wider body of knowledge?

- **Yes:** The theoretical framework/wider body of knowledge informing the study and the methods used is sufficiently described and justified.
- **Partial:** The theoretical framework/wider body of knowledge is not well described or justified; link to the study methods is not clear.
- **No:** Theoretical framework/wider body of knowledge is not discussed

### 5. Sampling strategy described, relevant and justified?

- **Yes:** The sampling strategy is clearly described and justified. The sample includes the full range of relevant, possible cases/settings (i.e., more than simple convenience sampling), permitting conceptual (rather than statistical) generalizations.
- **Partial:** The sampling strategy is not completely described, or is not fully justified. Or the sample does not include the full range of relevant, possible cases/settings (i.e., includes a convenience sample only).
- **No:** Sampling strategy is not described.

### 6. Data collection methods clearly described and systematic?

- **Yes:** The data collection procedures are systematic, and clearly described, permitting an “audit trail” such that the procedures could be replicated.
- **Partial:** Data collection procedures are not clearly described; difficult to determine if systematic or replicable.
- **No:** Data collection procedures are not described.

### 7. Data analysis clearly described, complete and systematic?

- **Yes:** Systematic analytic methods are clearly described, permitting an “audit trail” such that the procedures could be replicated. The iteration between the data and the explanations for the data (i.e., the theory) is clear – it is apparent how early, simple classifications evolved into more sophisticated coding structures which then evolved into clearly defined concepts/explanations for the data). Sufficient data is provided to allow the reader to judge whether the interpretation offered is adequately supported by the data.

- **Partial:** Analytic methods are not fully described. Or the iterative link between data and theory is not clear.
- **No:** The analytic methods are not described. Or it is not apparent that a link to theory informs the analysis.

#### 8. Use of verification procedure(s) to establish credibility of the study?

- **Yes:** One or more verification procedures were used to help establish credibility/trustworthiness of the study (e.g., prolonged engagement in the field, triangulation, peer review or debriefing, negative case analysis, member checks, external audits/inter-rater reliability, “batch” analysis).
- **No:** Verification procedure(s) not evident

#### 9. Conclusions supported by the results?

- **Yes:** Sufficient original evidence supports the conclusions. A link to theory informs any claims of generalizability.
- **Partial:** The conclusions are only partly supported by the data. Or claims of generalizability are not supported.
- **No:** The conclusions are not supported by the data. Or conclusions are absent.

#### 10. Reflexivity of the account?

- **Yes:** The researcher explicitly assessed the likely impact of their own personal characteristics (such as age, sex and professional status) and the methods used on the data obtained.
- **Partial:** Possible sources of influence on the data obtained were mentioned, but the likely impact of the influence or influences was not discussed.
- **No:** There is no evidence of reflexivity in the study report.

**Other relevant criteria**  
**by Irene Ng**

The following are four key criteria commonly used to assess the quality of research articles and academic thesis/dissertation. These are:

**Contribution**

- Does the article advance knowledge in/of the discipline?
- Are the findings and their implications noteworthy?
- Is the paper of interest to many people in the field or at least one segment of it(e.g., academics, practitioners, public policy makers, consumers etc.)?

**Communication**

- Is the article clearly written?
- Is the article laid out in a logical format?
- Are the major points easily grasped?