

Yale New Haven Health
Nursing Research and Evidence-Based Practice Committee
Research Literature Appraisal Tool

Article Number	Author(s): Article Title: Journal: Year Published: Volume: Number: Pages Numbers:
Level of Evidence and Grading: Fill in after completing appraisal (see Appendix A)	
Level of Evidence (Circle one): I II III IV V	Quality Grade (Circle one): High Good Low
Is this a reputable source of evidence? Yes <input type="checkbox"/> No <input type="checkbox"/>	

Appraisal Category		Summary	Appraisal
*Quantitative Study	#Qualitative Study		
Define independent & dependent variables	None used	Study purpose, aim, research questions and/or hypothesis:	Was information presented clearly? <input type="checkbox"/> Yes <input type="checkbox"/> No
Theoretical or conceptual framework	Philosophical underpinnings	Study framework or philosophical underpinnings, if evident:	Was information presented clearly? <input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
All relevant literature and or seminal work		Justification for the study: (problem statement [background] literature review)	Does this section address what is known and not known about the problem? <input type="checkbox"/> Yes <input type="checkbox"/> No Described how study would address gaps in knowledge? <input type="checkbox"/> Yes <input type="checkbox"/> No
See Appendix A Descriptive Quasi-experimental Experimental	See Appendix B Narrative Phenomenology Grounded theory Ethnography Case study	Study Methods: <i>Design</i>	Was design appropriate? <input type="checkbox"/> Yes <input type="checkbox"/> No

Appraisal Category		Summary	Appraisal
*Quantitative Study	#Qualitative Study		
No differentiation between study types		Study Methods: <i>Setting</i>	<p>Was the setting appropriate for study design?</p> <p><input type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>If multiple settings, were they appropriate for study design? <input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA</p>
<p>Probability sampling (i.e. random)</p> <p>Non-probability (i.e. convenience)</p> <p>Sample size: based on statistical test used and power analysis – goal to generalize results other populations</p>	<p>Purposeful or Theoretical sampling</p> <p>Sample size: based on judgment and experience often smaller than quantitative – goal to gain deeper understanding of concept</p>	Study Methods: <i>Sample (Describe sampling strategy, inclusion/exclusion criteria, sample size, and characteristics of sample [i.e. people, places, events])</i>	<p>Was sample size sufficient based on study design and data analysis?</p> <p><input type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>*Was sample representative of population under study?</p> <p><input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA</p> <p>*If an intervention was used were sample characteristics equivalent between control and intervention groups?</p>
<p>Data Collection Methods:</p> <p>Surveys (include response rate)</p> <p>Measurement instruments, tools, questionnaires)</p> <p>If intervention used, describe fidelity or how researcher made sure the intervention was consistently used with all subjects.</p>	<p>Data Collection Methods and techniques:</p> <p>Interviews, focus groups, observations, documents, (audio and videotaping, field notes)</p> <p>Collection and Analysis often occur simultaneously</p>	Study Methods: <i>Study Procedures (Describe *interventions, if tested, data collection methods, measurement instruments or data collection tools [including interview guides], timing/sequencing of data collection, human subjects protection)</i>	<p>Was data collection method described clearly? <input type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>Was data collection method a good fit with the study purpose and design?</p> <p><input type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>*For surveys, was response rate adequate (≥25% to 40%)?</p> <p>*Were measurement instruments validity and reliability discussed (psychometrically tested with adequate reliability (Chronbach alpha ≥0.70)?)</p> <p>*If intervention used, was it described clearly? <input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA</p> <p>#Was rigor discussed (credibility, transferability, dependability, confirmability) (see Appendix C Table 3)</p> <p><input type="checkbox"/> Yes <input type="checkbox"/> No</p>

Appraisal Category		Summary	Appraisal
*Quantitative Study	#Qualitative Study		
See Appendix C Descriptive statistics Bivariate analysis Multivariate analysis	See Appendix B Organizing data Reading & memoing Coding and themes Interpreting data Presenting data	Study Methods: <i>Data Analysis</i> (Describe methods used to analyze data)	Were the analysis methods appropriate? <input type="checkbox"/> Yes <input type="checkbox"/> No
No differentiation between study types		Results: (Summarize results)	Are results presented clearly? <input type="checkbox"/> Yes <input type="checkbox"/> No Are charts, graphs, tables easy to understand? <input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA If used, was description consistent with information found on them? <input type="checkbox"/> Yes <input type="checkbox"/> No *Were narratives used to support results? <input type="checkbox"/> Yes <input type="checkbox"/> No
No differentiation between study types		Limitations: (Summarize limitations)	Were limitations identified and addressed? <input type="checkbox"/> Yes <input type="checkbox"/> No
No differentiation between study types		Clinical Significance: (Focus on implications that this study has for nursing practice)	Does study contribute to nursing knowledge? <input type="checkbox"/> Yes <input type="checkbox"/> No Are the study results generalizable/transferrable to our practice setting? <input type="checkbox"/> Yes <input type="checkbox"/> No Do the results warrant examining our current practice for changes? <input type="checkbox"/> Yes <input type="checkbox"/> No

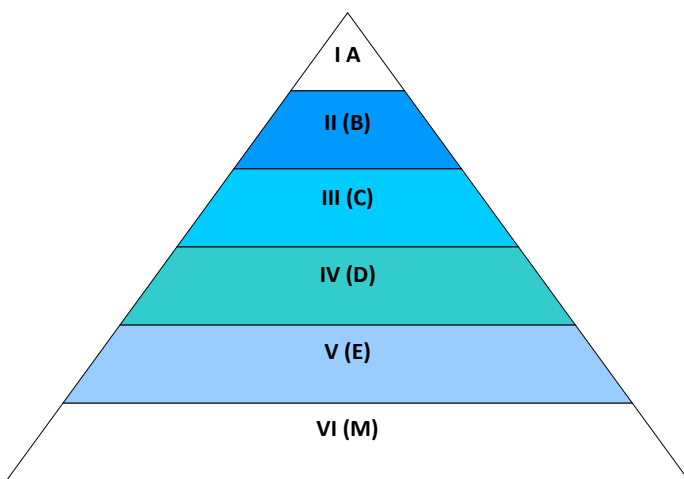
Appendix A

Level and Grading of Evidence by Project Methods

Level I Evidence	
Systematic Review	A summary of evidence, typically conducted by an expert or expert panel on a particular topic, that uses a rigorous process (to minimize bias) for identifying, appraising and synthesizing studies to answer a specific clinical question and draw conclusions about the data.
Meta-Analysis	A process of using quantitative methods to summarize the results from multiple studies obtained and critically reviewed using a rigorous process (to minimize bias) for identifying, appraising and synthesizing studies to answer a specific question and draw conclusions about the data gathered. The purpose of the process is to gain a summary studies (i.e. a measure of a single effect) that represents the effect of the intervention across multiple studies.
Randomized Controlled Trial (RCT)	A true experiment, (i.e., one that delivers an intervention or treatment), the strongest design to support cause and effect relationships, in which subjects are randomly assigned to control and experimental groups.
Level II Evidence	
Quasi-experiments	Design that test the effects of an intervention or treatment but lacks one or more characteristics of a true experiment (e.g. random assignment; control or comparison group)
Level III Evidence (Non Experimental)	
Cohort Study	Longitudinal study that begins with the gathering of two groups of patients (the cohort), one that received the exposure (e.g. to a disease) and one that does not, and then following these groups over time (prospective) to measure the development of different outcomes (diseases).
Case-Control Study	A type of research that retrospectively compares characteristics of an individual who has a certain condition (e.g., hypertension) with one who does not (i.e., a matched control or similar person without hypertension); often conducted for the purpose of identifying variables that might predict the condition (e.g., stressful lifestyle, sodium intake).
Cross Sectional Study	A study designed to observe an outcome or variable at a single point in time, usually for the purpose of inferring trends over time.
Correlational Descriptive Study	A study that is conducted for the purpose of describing the relationship between two or more variables.
Correlational Predictive Study	A study that is conducted for the purpose of describing what variables predicts a certain outcomes.
Descriptive Study	Studies conducted for the purpose of describing the characteristics of certain phenomena or selected variables.
Qualitative Study	Research that involves the collection of data in a nonnumeric form, such as personal interviews, usually with the intention of describing a phenomenon.
Level IV Evidence	
Clinical Practice Guidelines/ Consensus Panels	Opinion of respected authorities and/or nationally recognized expert committees/consensus panels based on scientific evidence i.e. National Guideline Clearinghouse

Level V Evidence (Based on experiential and non research evidence)	
Case Reports	Reports that describe the history of a single patient, or a small group of patients, usually in the form of a story.
Case Study	An intensive investigation of a case involving a person or small group of persons, an issue or an event.
Expert Opinion/ Manufacturer's Recommendations	

Melnyk, B. & Fineout-Overholt, E. (2011). *Evidence-based practice in nursing and healthcare: A guide to best practice (2nd Ed.)*. Philadelphia: Lippincott Williams and Wilkins.



Level of Evidence	Type of Evidence
Strongest I (A)	Evidence from systematic review or meta-analysis of multiple controlled studies with results that consistently support a specific action, intervention or treatment
II (B)	Evidence from at least one well designed controlled study, randomized & non-randomized, with results that support a specific action, intervention or treatment
III (C)	Evidence from qualitative studies, descriptive or correlational studies, integrative reviews or randomized controlled trials with inconsistent results
IV (D)	Evidence from peer reviewed professional organizational standards, with clinical evidence to support recommendations; Includes non-experimental studies
V (E) Weakest	Evidence from theory based evidence from expert opinion or multiple case reports; Interpretation of non-research based information by experts
VI (M)	Manufacturers' recommendations only

Based on: AACN's evidence-leveling system

Arnola, R.R., Bourgault, A.M., Halm, M.A., Board, R.M, Bucher, L, Harrington, L., Heafey, C... & Medina, J. (2009). Upgrading the American Association of Critical-Care Nurses' evidence-leveling hierarchy. *American Journal of Critical Care*, 18, 405-409.

Level of Evidence	Quality Grading Guides
Level I	A High quality: consistent results, sufficient sample size, adequate control, and definitive conclusions; consistent recommendations based on extensive literature review that includes thoughtful reference to scientific evidence.
Level II	B Good quality: reasonably consistent results, sufficient sample size, some control, and fairly definitive conclusions; reasonably consistent recommendations based on fairly comprehensive literature review that includes some reference to scientific evidence
Level III	C Low quality or major flaws: little evidence with inconsistent results, insufficient sample size, conclusions cannot be drawn.
Level IV	<p>A High quality: well-defined, reproducible search strategies; consistent results with sufficient numbers of well-designed studies; criteria-based evaluation of overall scientific strength and quality of included studies, and definitive conclusions</p> <p>B Good quality: reasonably thorough and appropriate search; reasonably consistent results, sufficient numbers of well-designed studies, evaluation of strengths and limitations of included studies, with fairly definitive results</p> <p>C Low quality or major flaws: undefined, poorly defined, or limited search strategies; insufficient evidence with inconsistent results, conclusions cannot be drawn</p>
Level V	<p>A High quality: expertise is clearly evident.</p> <p>B Good quality: expertise appears to be credible.</p> <p>C Low quality or major flaws: expertise is not discernable or is dubious.</p>

Appendix B

Table 1: Traditions of Qualitative Research (Study Methods)

Tradition	Purpose	Key Elements
Narrative	Exploring the life of a single individual or small group of individuals	<ul style="list-style-type: none"> • Studies one or more individuals • Uses interviews primarily • Develops narratives, usually chronologically, about lives
Phenomenology	Understanding the lived experience of a phenomenology	<ul style="list-style-type: none"> • Studies multiple people experiencing the same phenomenon • Uses interviews primarily • Uses data saturation for sampling • Describes the “essence” of the experience that is shared
Grounded Theory	Developing theory based on field-collected data	<ul style="list-style-type: none"> • Studies a process or action • Uses interviews primary • Uses open, axial, and selective coding • Uses theoretical sampling • Generates a graphical representation of the theory
Ethnography	Describing elements of a culture-sharing group	<ul style="list-style-type: none"> • Studies a group with the same culture • Uses observations and interviews • Analyzes data to determine cultural traits shared by a group
Case Study	Developing an understanding of a single case or multiple related cases	<ul style="list-style-type: none"> • Studies an event or activity, or multiple persons • Analyzes cases to determine themes within and between cases

Source: Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Table 2: Data Analysis in Qualitative Research

Data Analysis Step	Details
Organizing Data	Converting raw data into organized units such as transcribed interviews into electronic format
Reading and Memoing	Reviewing the entirety of data collected for immersion before development of codes and themes
Coding and Developing Themes	Categorizing pieces of data into codes (small categorizes of information) and reducing codes into themes (broad units of categories comprised of codes)
Interpreting Data	Drawing connections between themes and codes to view a larger picture of the concept being studied
Presenting the Data	Using graphical, tabular, or text format to present the interpretation of data

Source: Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Table 3: Methodological Rigor in Qualitative Research

Element	Description
Credibility	The degree to which the data collected are accurate, for example through member checking, triangulation, and negative case analysis
Transferability	The degree to which the findings can be transferred to another group of individuals (rather than generalized to an entire population)
Dependability	The degree to which the steps of the qualitative research process are described within the manuscript and the steps are “transparent”
Confirmability	The degree to which the researcher’s experiences and mindset to the concept are integrated into the data collected and conclusions reached.

Source: Tappen, R. M. (2011). *Advanced nursing research: From theory to practice*. New York: Jones and Bartlett Publishing.

11/29/16

Appendix C

Choosing the Appropriate Statistical Test: Marge Funk, PhD, RN

Bivariate Statistical Tests

Test Name	Independent or Related	Purpose	Measurement Level	
			IV	DV
Parametric Tests				
Independent t-test	I	Test the difference between 2 independent group means	N	I/R
Paired t-test	R	Test the difference between 2 related group means	N	I/R
1-way analysis of variance (ANOVA)	I	Test the difference among the means of 3+ independent groups	N	I/R
Repeated measures ANOVA	R	Test the difference among the means of 3+ related groups or sets of scores	N	I/R
Pearson correlation	I, R	Test the existence of a relationship between 2 variables	I/R	I/R
Linear regression	--	Predict value of DV for given value of IV	I/R	I/R
Nonparametric Tests				
Mann-Whitney U-test	I	Test the difference in ranks of scores of 2 independent groups	N	O
Wilcoxon signed-rank test	R	Test the difference in ranks of scores of 2 related groups	N	O
Kruskal-Wallis test	I	Test the difference in ranks of scores of 3+ independent groups	N	O
Friedman test	R	Test the difference in ranks of scores of 3+ related groups	N	O
Chi square test	I	Test the difference in proportions in 2+ independent groups	N	N
McNemar test	R	Test the difference in proportions for 2 related groups (2x2)	N	N
Cochran's Q test	R	Test the difference in proportions for 3+ related groups	N	N
Fisher's exact test	I	Test the difference in proportions in 2 independent groups when N < 30, any expected cell frequency < 5, or cell with observed frequency of 0	N	N
Phi coefficient or odds ratio	I	Examine the magnitude of a relationship between 2 dichotomous variables	N	N
Cramer's V	I	Examine the magnitude of a relationship between 2 variables (not restricted to dichotomous)	N	N
Spearman's rho	I, R	Test the existence of relationship between 2 variables	O	O

IV, Independent variable; DV, dependent variable; I, independent; R, related; N, nominal; O, ordinal or **non-normally distributed interval/ratio**; I/R, interval/ratio.

Note: On some tests, the measurement level of the IV & DV can be switched.

Multivariate/Multivariable & Advanced Statistical Tests

1. ANOVA

a. One-way ANOVA (bivariate)

- Purpose: Test the difference among the means of ≥ 3 groups.
- Variables: IV = 1 N; DV = 1 I/R

b. Repeated measures ANOVA (bivariate)

- Purpose: 1) Repeated measures (≥ 3) of DV on same subjects over time; 2) Exposure of all subjects to all treatment conditions (≥ 3).
- Variables: IV = 1 N; DV = 1 I/R

c. Two-way ANOVA

- Purpose: Test main effect of each IV on DV and test interaction between 2 IVs.
- Variables: IV = 2 N; DV = 1 I/R

d. ANCOVA

- Purpose: Test effect of IV on DV while controlling for covariate(s).
- Variables: IV = 1 N; DV = 1 I/R; Covar = ≥ 1 I/R (sometimes N)

e. Mixed-Design ANOVA

- Purpose: Extension of repeated measures ANOVA but with ≥ 2 groups
- Variables: IV = ≥ 2 N (1 is usually time); DV = 1 I/R

f. MANOVA

- Purpose: Test the difference among the means of ≥ 2 groups for ≥ 2 DVs simultaneously.
- Variables: IV ≥ 1 N; DV ≥ 2 I/R

2. Regression

a. Simple linear regression (bivariate)

- Purpose: 1) Determine if a linear relationship exists between IV and DV; 2) Predict value of DV based on given value of IV.
- Variables: IV = 1 I/R; DV = 1 I/R

b. Multiple regression

- Purpose: 1) Test the relationship between 2+ IVs and 1 DV; 2) Determine if an IV is r/t the DV in the presence of or accounting for other factors; 3) Predict value of DV based on several IVs; 4) Determine the amount of variability in DV that is explained by IVs.
- Variables: IV > 1 any level; DV = 1 I/R

c. Logistic regression

- Purpose: 1) Test the relationship between 2+ IVs and 1 DV; 2) Determine if an IV is r/t the DV in the presence of or accounting for other factors; 3) Determine predictors of a particular outcome.
- Variables: IV > 1 any level; DV = 1 N (dichotomous)

3. Survival Analysis (e.g., life table or actuarial analysis; Kaplan-Meier method; log-rank test; Cox proportional hazard model)

- Purpose: Determine time to an endpoint when subjects enter study at different times and some subjects may not have reached the endpoint at end of data collection.
- Variables: N/A

4. Measurement Statistics

a. Evaluation of agreement

- Cohen's Kappa: nominal or ordinal
- Intraclass correlation coefficient: interval/ratio

b. Evaluation of consistency

- Cronbach's alpha

c. Comparison of methods

- Bland-Altman: interval/ratio measured on same scale

Steps to Determine Appropriate Test to Use

1. Identify variables (IV vs. DV – be aware of sample)
2. Measurement level of the variables (nominal, ordinal, interval/ratio)
3. # of groups being compared (for nominal variables)
4. Whether the groups are independent or related (measured in same people over time; matched)
5. Whether the dependent variable is normally distributed (use parametric vs. nonparametric test)
6. Sample size
7. # of variables (use univariate, bivariate, or multivariate statistics)
8. If >2 variables . . .
 - a. Determine IV(s) and DV(s) and their level of measurement
 - b. Determine purpose, e.g.
 - a. Interaction
 - b. Involve repeated-measures factors & between-group factors
 - c. Prediction
 - d. Association of IV(s) with DV in presence of other factors
 - e. Amount of variability in DV explained by IVs
 - f. Time to endpoint

approp test handout ynhh 1-5-17