

CRITERIA FOR UPDATED TOOL EVALUATION

Following are criteria and associated indicators used in the evaluation of tools for assessing pain in nonverbal elders. The indicators are based on measurement theory and geriatric pain literature. Sources are noted at the end of the document. Each tool was rated for evidence that supported the criteria and indicators as defined using a 4-point scale illustrated at the end of the document.

Criteria	Indicators
Conceptualization	<p>Purpose:</p> <ul style="list-style-type: none"> • The purpose of the instrument is clearly defined. • The purpose is appropriate for non-verbal elders. <p>Conceptual basis:</p> <ul style="list-style-type: none"> • The concept (pain) is clearly defined and is used consistently throughout. • There is no blurring with competing concepts. <p>Item generation:</p> <ul style="list-style-type: none"> • The method of item generation is described and appropriate (eg. clinical observation, focus groups, literature review). • Tool items address pain behaviors identified in literature and evidence based guidelines (eg. AGS, 2002). • Tool items are a representative and comprehensive sample of the content area. • The scoring procedure is conceptually sound. <p>Face validity:</p> <ul style="list-style-type: none"> • The tool appears to measure pain in non-verbal elders. • The tool appears relevant to the stated purpose: to measure pain in non-verbal elders. <p>Content validity:</p> <ul style="list-style-type: none"> • Method of evaluating content validity is clearly defined. • The tool has been evaluated through an external process by independent reviewers with expertise in pain in elders with dementia. • Results of content validity evaluation are within acceptable range. • Language/cultural equivalence is reported or described for tool use in English.
Subjects	<p>Focus on acute or long term care setting is identified.</p> <p>Intended subjects are identified as being non-verbal or non-communicative.</p> <p>Identification of subjects' cognitive impairment included standard assessment approaches.</p> <p>The instrument is tested in a group of the intended subjects.</p> <p>The tool has good fit to the demographic and cultural background of the intended subjects.</p> <p>Characteristics of the sample group are provided.</p> <p>Sample size is adequate for tool testing. (Rule of thumb: 5 subjects per tool item).</p>
Administration, Scoring and Feasibility	<p>Method of administration is described.</p> <p>Scoring procedures are clearly described.</p> <p>Interpretation of tool score is clear.</p> <p>There is consistency in format across items.</p> <p>Clinical utility, including time, training and skill needed, is reported.</p>
Reliability	<p>Internal consistency reliability:</p> <ul style="list-style-type: none"> • Method of internal consistency reliability testing is appropriate for the data. • Characteristics of the raters are appropriate. • Correlation coefficient is within acceptable levels.

	<p>Interrater (or interobserver) reliability:</p> <ul style="list-style-type: none"> • The measure of interrater reliability is appropriate for type of data the tool produces: Pearson's r for interval data, percent agreement or kappa for nominal and ordinal data. • Characteristics of raters are appropriate. • Correlation coefficient is within acceptable levels. <p>Test-retest reliability:</p> <ul style="list-style-type: none"> • Scores from the same subject measured at two different times are correlated. • The interval between tests is provided and appropriate. • Characteristics of raters are appropriate. • Correlation coefficient is within acceptable levels. <p>Intrarater reliability:</p> <ul style="list-style-type: none"> • The event being measured by the rater is defined. • Time 1 and time 2 are clearly defined. • Correlation coefficient is within acceptable levels. • Characteristics of raters are appropriate.
Validity	
	<p>Criterion related validity (eg. predictive validity, concurrent validity) or construct validity (eg. known groups method):</p> <ul style="list-style-type: none"> • The type of validity testing conducted is appropriate. • Gold standards for comparison are valid for the population. • Sufficient detail regarding procedures is provided. • Analysis techniques are appropriate for data and type of validity procedure.

Key to evaluation criteria

Conceptualization is a composite score based on purpose, conceptual basis, item generation and content validity

Reliability is a composite score including test-retest and/or intrarater reliability, internal consistency and interrater reliability

Validity is a composite score including criterion related and/or construct validity

Key to rating criteria and indicators

3= Available evidence is strong

2= Available evidence supports need for further testing

1= Available evidence is insufficient and/or tool revisions are needed

0= Evidence is absent

Key to categories for psychometric statistics used for this review

<p><u>Cronbach's alpha</u> no info (0) $\alpha < .60$ inadequate. (1) $\alpha = .60- .70, > .90$ fair. (2) $\alpha = .70 - .90$ sufficient. (3)</p>	<p><u>Cohen's Kappa</u> no info (0) $\kappa < .60$ = inadequate (1) $\kappa .61- .80$ = good (2) $\kappa .81-1.00$ = very good (3)</p>	<p><u>Cohen's r</u> no info (0) $r = .10- .29$ = small; (1) $r = .30- .49$ = medium; (2) $r = \text{or} > .5$ large. (3)</p>	<p><u>Intra-class correlation Coefficient r</u> not rated</p>
--	--	--	---

Sources:

1. Jacobson, S.F.(1997). Evaluating instruments for use in clinical nursing research. In M. Frank-Stromborg & S.J. Olsen (eds.). Instruments for Clinical Health Care Research. Boston, MA: Jones and Bartlett Publishers.
2. Norbeck, J.S. (1985). What constitutes a publishable report of instrument development? *Nursing Research*, 34(6), 380-382.
3. Burns, N., & Grove, S. K. (2001). *The Practice of Nursing Research: Conduct, Critique, & Utilization* (4th ed.). Philadelphia: W.B. Saunders Company.
4. AGS Panel on Persistent Pain in Older Persons, (2002). The Management of persistent pain in older persons. *Journal of the American Geriatric Society*, 50, S205-S224.

Evaluation criteria developed by: K. Herr, S. Decker, K. Bjoro, H. Bursch, The University of Iowa.

Contact information: keela-herr@uiowa.edu

Revised 05/08