Tool: The Abbey Pain Scale

Tool developer: Abbey, J.A., DeBellis, A, Piller, N., Esterman, A., Parker, D., Giles, L. & Lowcay, B.

Country of origin: Australia

Conceptualization Panel rating: 1	
Purpose	The stated purpose was to develop a pain scale for people with late-stage dementia that was efficient, effective and able to be used by a variety of care staff.
Conceptual basis	There is no discussion or presentation of the conceptual basis for the tool. The tool attempts to measure severity of "acute pain", "chronic pain" and "acute on chronic." Yet a clear discussion of the conceptual basis for the tool is absent.
Item Generation	Tool items • Vocalization • Facial expression • Change in body language • Behavioral change • Physiological change • Physical change Each item is leveled on a four point scale for severity (Absent: 0; Mild: 1; Moderate: 2; Severe 3). Individual item scores are summed to arrive at a total score ranging from 0-18. The total score is interpreted as follows: No pain: 0-2; Mild: 3-7; Moderate: 8-13; Severe: 14+. Item generation process Phase I: A draft of the scale was based on: • Literature review • A Delphi study by geriatric and pain experts • Focus group of practitioners. Phase II: A prototype of the tool consisting of 12 items was generated in which each of the 6 items listed above was measured on both a dichotomous scale and an ordinal scale. Regression analysis was used to determine which of the twelve items were most predictive of the de facto gold standard (nurses' holistic impression of pain severity). Twelve items predicted 42% of the variability on the holistic pain question. Items measuring pain severity were more powerful predictors of holistic pain than presence/absence of pain. Thus, items rated as present/absent were deleted. Six severity items were retained accounting for 41% of the variability in the scores.
Content Validity	The draft scale was reviewed by gerontological and pain experts through a Delphi study before being tested clinically and also by discussion with practitioners using focus group debate.
-Panel Commentary	There is conceptual blurring between acute and chronic pain. There is no discussion in the paper on characteristics of the one type of pain in relationship to the other or the overlap (e.g. physiologic changes vs. behavioral change). There is conceptual blurring between pain behaviors and pain etiology, (e.g.

	presence of skin tears or arthritis), both elements of the rating scale.
	Development of the tool is based on the assumption that caregivers can reliably rate the intensity of pain in elders, although the interpretation of pain severity has not been substantiated in the research on pain in elders with dementia. Nurses' holistic impression of pain severity was used as the gold standard. However, this assumption is not supported by current literature, which indicates that caregivers can detect presence of pain, but not severity or intensity. Therefore, using "the gold standard of caregiver's report of pain intensity" in regression analysis to derive tool items is a limitation. Little information is available on the Delphi study: who the experts were, what items were presented in the Delphi and the results. Also, almost no information is available on the focus group: who were the participants, and their expertise (educational and clinical background and relevance to elders with dementia and pain). Comprehensiveness and clarity of items The tool includes at least one cue from each of the 6 categories of non-verbal pain behaviors in the AGS Persistent Pain Guidelines. Facial expression, Verbalizations/vocalizations, Body language and to some extent Changes in activity patterns or routines, Mental status changes and Changes in interpersonal interactions. However, inclusion of physiological indicators may compromise the overall tool scoring of chronic pain. The rationale for including pain etiology items in a tool for assessing pain that totals overall severity of behaviors in non-verbal elders is not clear. The tool is intended to measure acute pain, chronic pain and acute on chronic pain. However, it is unclear how the tool could differentiate pain types. Moreover, the tool has not been formally evaluated for content validity.
Subjects **Panel rating: 1**	
Subjects	The tool development and testing took place in 24 aged care residential facilities in four Australian states: South Australia, New South Wales, Queensland and Victoria. Inclusion criteria: Residents with late-stage dementia as confirmed by a senior registered nurse and who were perceived by facility staff as experiencing pain during the project. Staff observations were made by 61 staff who completed the pain scale: 45 registered nurses and 7 enrolled nurses. In the initial stage of tool testing, data were available for 52 residents for a total of 770 pain episodes (non-independent scores). In stage II of tool testing, data were available for 61 residents for a total of 236 pain episodes. Age: 83 years (median), Range: 60-97 years. Gender: Female: 66%, Male 34%.

-Panel Commentary	Focus on long term care setting is clearly identified. However, the study was carried out at 24 sites with 61 staff members with varied skill level completing observations. This allows for wide variation and consequently potential for measurement error.
	A standard assessment approach to document presence of dementia is not included. Relying on senior nurses' confirmation of dementia allows inconsistent assessment across study sites.
	The study subjects are limited to those who were identified by the staff as having pain. It is not known whether pain assessment was carried out on a regular basis. Patients with atypical presentation of pain may not have been identified for participation in the study.
	Age of subjects is appropriate. Gender: Female: male ratio is adequate. There is no information on ethnic/racial diversity.
	It is unclear if the samples are independent**.
	Using 5 subjects per tool item as a rule of thumb, a minimum sample size of 60 subjects (12 items x 5 subjects) and 30 subjects (6 items x 5 subjects) in stage 1 and stage II respectively would be needed. Thus, with 52 subjects in stage I this sample is not sufficient for regression analysis; however, with 61 subjects in stage II the sample is sufficient for tool evaluation.

Administration, Scoring, Feasibility
Panel rating: 1

Administration, Scoring, Feasibility	Instructions for using the Abbey Pain Scale are presented on a poster. Nurses are asked to use the tool when pain is suspected. Prompts depicted on the poster include "These people have dementia." "Are they in pain?" "If they can't tell you, use the one-minute Abbey Pain Scale."
	There are a few limited instructions on the tool schema. Scoring prodecure is described above.
	The rater is asked to indicate what type of pain the subject has: acute, chronic or acute on chronic.
	There are qualitative reports that the tool took less than one minute to complete.
-Panel Commentary	Method of administration and scoring is not adequately described. • Nurses are asked to use the tool when pain is suspected. It is unclear what triggers the pain assessment. A lack of systematic pain assessment may result in non-detection of pain. Systematic pain assessment may have been more appropriate in a tool development study. This gives rise to the issue of scheduled pain assessment vs. prn assessment. • Scoring procedures are unclear. There are no instructions on what constitutes the different severity levels for each item. • There is no justification for the total severity scoring system or its interpretation. • There is no indication as to how the rater arrives at "type of pain" or how this impacts treatment decisions. Clinical utility • Time: The tool developers report the tool takes less than one minute to

	score. However, no data are provided.
	• <u>Skill needed</u>: The tool developers do not specify skill level needed. All nursing personnel are referred to as "staff." It is unclear if there are any limitations to the scope of practice of those using the tool.

Reliability **Panel rating: 1**	
Internal consistency	Internal consistency was assessed in stage II of tool testing. Data were available for 61 residents for a total of 236 pain episodes (see sample characteristics above). Internal consistency reliability reported: Pre-intervention: α=.74 and post-intervention: α=.74.
Interrater reliability	Interrater reliability was assessed in stage II of the testing. Residents (n=18) were assessed by two staff members. Characteristics of the wider sample of subjects are provided above. Assessed by intra-class correlation coefficient: Pre intervention: ICC=0.63 (p=0.02) Post-intervention: ICC=0.44 (p=0.12).
Test-retest reliability	No test retest or intrarater reliability is reported.
-Panel commentary	<u>Internal consistency</u> Although the tool has an acceptable level of internal consistency, it is unclear what data were used for analysis, eg. patients, mean pain score, pain episodes. <u>Interrater reliability</u> Staff qualifications are not specified. Intraclass correlation is an appropriate test for interrater reliability. However, the reliabilities are low, particularly for post intervention evaluation. <u>Test-retest reliability</u> No test-retest or intrarater reliability is available. Thus, tool reliability has not been adequately established based on available information.

Validity: Criterion or construct **Panel rating: 1**	
Construct validity/ Criterion related validity	Validity testing was conducted in stage II of testing. Sample included 61 residents (see sample characteristics above) with a total of 236 pain episodes. The number and types of intervention varied from patient to patient. However, the most common intervention was analgesics, followed by repositioning. Staff (n=61) completed the pain scale: 45 (74%) were registered nurses and 7 (12%) were enrolled nurses. <u>Concurrent validity</u> The pre-intervention pain score was associated with the holistic impression of pain as assessed by the nurse. The holistic impression was rated on a scale of 1=no pain, 2=mild, 3=discomforting, 4=distressing or 5=severe. Results: Gamma: 0.586 (p<.001). <u>Predictive validity</u> Change in pain score before and after intervention:

	Average pain score (n=61): Pre-intervention (n=61): 9.02 ($\pm$ 3.75) Post-intervention (n=61): 4.21 ($\pm$ 3.20) Paired t-test: showed the reduction to be highly statistically significant (p<0.001).
-Panel commentary	Concurrent validity The holistic impression of pain as assessed by the nurse was used as a gold standard. However, the expertise or competence to assess pain in persons with dementia is not reported. There is no indication as to whether the Abbey Pain Scale scores and the holistic impression of pain are independent measures or assessed by the same nurse, which would impact data validity. Validity data is based on 61 different staff using the tool which presents bias for evaluation and interpretation. Gamma is an appropriate measure of association for measures at ordinal level or higher. Gamma varies between -1 and 1. The results in this study indicate that the two measures are moderately positively associated. Predictive validity Variability in interventions and their effectiveness limits ability to make conclusions regarding change in pain. However, because data reported are in the expected direction regarding change in pain score, future testing in controlled circumstances is warranted.

Summary of panel evaluation of pain assessment tool

The Abbey Pain Scale includes key behaviors representative of the scope of behavioral pain indicators in persons with dementia. However, the tool also includes items such as physiological changes and pain etiologies, which are not behavioral pain indicators and are not conceptually congruent with the intent of the tool. Moreover, the ability of health care providers to determine severity of pain from behavioral indicators has not been established. The subject sample is appropriate, however validation in minority samples and use of a standard assessment of cognitive impairment are needed. Scoring procedures and training to use the tool are not clear, nor is qualification of those individuals providing the gold standard. Tool reliability is not supported with current data available. Validity testing based on nurse judgment of pain severity is not substantiated in the literature, particularly without evidence supporting the expertise of the raters. Tool revision and additional testing in controlled circumstances are recommended.

Source of evidence

Abbey, J.A., Piller, N., DeBellis, A, Esterman, A., Parker, D., Giles, L. & Lowcay, B. (2004). The Abbey Pain Scale. A 1-minute numerical indicator for people with late-stage dementia. International Journal of Palliative Nursing, 10(1), 6-13.

Key to panel rating

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

Evaluation completed by:

K. Herr, S. Decker, K. Bjoro, University of Iowa.
Contact information: keela-herr@uiowa.edu