Nursing Research • January/February 2007 • Vol 56, No 1, 34–43

Observation Scales for Pain Assessment in Older Adults With Cognitive Impairments or Communication Difficulties

Rhodee van Herk ▼ Monique van Dijk ▼ Frans P.M. Baar ▼ Dick Tibboel ▼ Rianne de Wit

- ***Background:*** Several pain observation scales have been developed to accurately assess and manage pain in older adults with severe cognitive impairments, communication difficulties, or both.
- ***Objective:*** To review relevant pain observation scales and the psychometric qualities of these scales.
- ***Methods:*** The literature was searched for articles reporting the use of a pain observation scale in an empirical study and describing psychometric properties in older adults with cognitive impairments, communication difficulties, or both.
- ***Results:*** Thirteen pain observation scales were included. Scales differed in numbers of items, types of categories, and psychometric properties. Facial expression, vocalization, motor behavior, and social behavior or mood are categories present in most of the scales. In terms of reliability and validity, however, most studies are too limited or incomplete to allow definite conclusions to be drawn about usefulness in daily practice.
- ***Discussion:*** As different methods of evaluating reliability and validity were used, and different aims (e.g., type of pain) were pursued, the available scales cannot be compared easily. Nevertheless, a few are promising, given preliminary results. These should be examined further on psychometric properties and usefulness in different populations because optimal pain assessment is necessary for efficient and effective pain treatment.

- ***Key Words:*** cognitive impairment · communication difficulties · older adults · pain measurement

Worldwide, the proportion of people older than 60 years is growing faster than any other group. The World Health Organization (WHO) expects that the group will have grown by 223% between 1970 and 2025 (WHO, 2002). The consequences of aging are important for society, in general, and for healthcare facilities and organizations, in particular. A specific problem is pain, which is a serious and often unrecognized health problem for older adults. The reported prevalence of pain in older adults varies, and may reach 83% in nursing homes (Ferrell, Ferrell, & Rivera, 1995).

Pain is not always assessed and managed adequately. Because pain is a subjective experience, self-report is usually considered the gold standard. Although tools such as numeric rating scales are appropriate for use in older adults with mild to moderate cognitive impairments, they may be of little help in persons with severe or advanced cognitive impairments (Closs, Barr, Briggs, Cash, & Seers, 2004). Self-reports may be biased or impossible in persons who are cognitively impaired or have limited communicative skills. Therefore, these persons are at even higher risk for undertreatment of pain. Several studies reported that persons with communication difficulties received less pain medication than did verbal persons (Morrison & Siu, 2000).

Thus, pain assessment in persons with severe cognitive impairments, communication difficulties, or both should include observations of behavior. In 2002, the American Geriatrics Society Panel on Persistent Pain in Older Persons published six common pain indicators: facial expressions, verbalizations or vocalizations, body movements, changes in interpersonal interactions, changes in activity patterns or routines, and mental status changes. Closs, Cash, Barr, and Briggs (2005) identified pain behavior in nursing home residents with different levels of cognitive impairment. They found three main groups of cues: verbal and body language cues, acute behavioral cues, and general changes in behavior or mood. Body movements were seen most frequently in persons with severe cognitive impairments (Closs et al., 2005). Facial activity provides the most sensitive and specific nonverbal response during a painful event

Rhodee van Herk, MA, *is PhD student, Pain Expertise Center, Erasmus MC, Rotterdam, The Netherlands.*

Monique van Dijk, PhD, *is Senior Researcher, Quality of Care; and* ***Dick Tibboel, PhD,*** *is Professor, Department of Pediatric Surgery, Erasmus MC–Sophia Children's Hospital, Rotterdam, The Netherlands.*

Frans P.M. Baar, MD, *is Head of Nursing Home, Laurens, Antonius IJsselmonde, Rotterdam, The Netherlands.*

Rianne de Wit, PhD, *is Professor, Department of Nursing Sciences, Maastricht University en University Hospital Maastricht, The Netherlands.*

(Craig, Prkachin, & Grunau, 1992), and facial activity can be a reliable parameter to assess pain in persons with communication difficulties (Manfredi, Breuer, Meier, & Libow, 2003). Manfredi et al. (2003) studied the reliability and validity of facial expressions as pain indicators in persons with severe dementia. They concluded that clinicians' observations of facial expressions and vocalizations are accurate means for assessing the presence of pain (not intensity) in persons with severe or advanced dementia. Some observation scales include physiological items to assess pain. There is little evidence, however, that specific physiological indicators would be strong indicators of chronic pain.

The purpose of this study was to review the pain observation scales used in or developed for older adults with severe cognitive impairments, communication difficulties, or both. Attention is paid to the different evaluation criteria and psychometric properties of these scales; furthermore, guidelines are proposed for the development of the ideal assessment scale.

Methods

Literature Search

Four computerized bibliographic databases (PubMed Medline, PsycINFO, Cinahl, and PiCarta) were screened for publications from 1980 through 2005. MeSH headings used were: pain (measurement) AND dementia OR Alzheimer OR (aged/aged, 80 and over/frail elderly) AND communication disorders. The reference lists of retrieved articles were also searched for additional references.

Articles were included if a pain observation scale was used in an empirical study, and psychometric properties were reported in older adults with severe cognitive impairments, communication difficulties, or both, or if a scale had been developed specifically for use in older adults. Pain observation scales developed specifically for children and critically ill, sedated hospitalized patients were excluded. Also, non-English articles were excluded, unless an English abstract was available.

Criteria

New instruments need to be tested for their psychometric properties before they can be used in practice. The instruments in this review were therefore assessed on relevant criteria, as described below.

Reliability

Interrater reliability is the agreement between two or more raters when both are rating the same person. Intrarater reliability is the consistency of scores assigned by the same rater at different times. Cohen's Kappa or intraclass correlation coefficients can be used to express degree of reliability. Kappa < .20 is considered poor; .21–.40, fair; .41–.60, moderate; .61–.80, good; and .81–1.00, very good (Altman, 1991; Landis & Koch, 1977). The higher the coefficient, the more reliable the agreement between raters.

Test–retest reliability is the extent to which a stable condition tends to produce similar scores over time. Because pain does not necessarily remain stable from one day to the next, test–retest stability coefficients have limited usefulness as estimates of the reliability of pain scales (Jensen, 2003). This outcome is left out of consideration in the tables.

The basic assumption of internal consistency is that all items of a scale address the same theoretical construct. Thus, a scale is considered to be internally consistent when there is a high intercorrelation among the scores of the items. A Cronbach's alpha coefficient (α) of .90 or higher indicates high internal consistency, but also redundancy among items. A coefficient of magnitude between .70 and .90 is adequate for group-level comparisons. Scales with Cronbach's alpha lower than .70 are inadequate for most purposes (Nunnally & Bernstein, 1994).

Validity

Generally considered the most important metric property of a scale, validity is the degree to which a scale measures what it is supposed to measure (Polit & Hungler, 1987).

Face validity is the extent to which the test or procedure on first impression appears to measure what it is intended to measure.

Content validity is the extent to which the items of a scale are representative of some defined universe or domain of interest. Content validity is usually determined by expert judgment (Jensen, 2003). Overall, face and content validity have been well established for most pain scales and, when proven satisfactory, only indicate beginning validity. Therefore, these will be left out of consideration in this review.

Criterion validity assesses the relationship of a scale and a particular criterion. One aspect of criterion validity is predictive validity: the extent to which a scale is able to predict important outcomes (Jensen, 2003). Another aspect of criterion validity is concurrent validity, which refers to the comparability of a scale with a criterion. Because there is no gold standard for a nonverbal population, concurrent (or congruent) validity is usually determined by correlating a pain observation score with proxy reports or other existing pain observation scales.

Construct validity is the extent to which a scale assesses the specific domain or construct of interest. The most common sources of construct validity are the associations, often expressed by correlation coefficients, between scales of the same construct (i.e., pain) using different methods (convergent validity), or between scales of different constructs (i.e., pain and fear) or groups known to have a large amount of the construct versus those that do not, using the same method (discriminant validity). Cohen's criteria were used to judge the value of correlation coefficients: .10 to .29 (small r); .30 to .49 (medium r); and $\geq$.50 (large r; Cohen, 1988).

Sensitivity to change or responsiveness of a scale has been mentioned as an aspect of validity. Using a sensitive or responsive pain scale, it should be possible to detect changes, and the scale results will remain stable when no change has occurred.

Feasibility and Clinical Utility

The feasibility of a scale is its applicability in daily practice, especially regarding its ease of use and the length of time to complete it. Clinical utility refers to the usefulness of the measure for decision making. This may be established by

calculating cutoff scores that discriminate between pain and no pain. Cutoff scores enable pain assessment to be coupled with a treatment algorithm.

Results

In the literature search, 13 pain observation scales were found that met the requirements for this study. The structural characteristics are presented in Table 1, and the major psychometric properties are presented in Tables 2 and 3.

Structural Characteristics

One of the oldest observational scales, the Facial Action Coding System (FACS) enables to study different emotions, including pain (Ekman & Friesen, 1978). Explicit criteria are employed using the FACS to identify 46 discrete facial action units (AUs) involving specific muscles or groups of muscles, such as brow raise, brow lower, upper lip raise, lip stretch, or mouth stretch. Coders note frequency and intensity of facial AUs. The FACS was used in several studies of pain in cognitively impaired older adults (Hadjistavropoulos, Craig, Martin, Hadjistavropoulos, & McMurtry, 1997; Hadjistavropoulos, LaChapelle, Hadjistavropoulos, Green, & Asmundson, 2002; Hadjistavropoulos, LaChapelle, MacLeod, Snider, & Craig, 2000; LaChapelle, Hadjistavropoulos, & Craig, 1999). Pain stimuli were acute phasic pain by injection and pain experienced on exercise after surgery. Both interrater and intrarater reliability were found to be moderate to high. Significant correlations were found between FACS scores and self-reported pain intensity, and small to medium correlations were found between FACS and another pain observation scale (Pain Behavior Method [PBM]; Keefe & Block, 1982) for some items. In all studies, significant differences were found between painful and painless situations, but no differences in pain scores were found between persons with and without analgesics. Overall, the FACS seems to be a reliable and valid tool, and extended information on its use is available via a Web site (http://face-and-emotion.com/dataface/facs/description.jsp). However, the unidimensional focus on facial expressions seems disadvantageous, and the complexity of learning the system and the use of video recordings make the FACS unsuitable for clinical practice.

The PBM (Keefe & Block, 1982) was developed to measure pain in persons with chronic low back pain. In this observation system, five pain behaviors (guarding, bracing, rubbing, grimacing, and sighing) are examined during walking, shifting, sitting, standing, and declining. Participants are videotaped for 10 minutes, and recordings are later scored for the presence or absence of each behavior. Hadjistavropoulos et al. (2000) studied the psychometric

TABLE 1. Structural Characteristics

Pain Observation Scale	Target Population	No. of Items	Categories*	Response Category	Score Range
Facial Activity Coding System (FACS)	Cognitively impaired (older) adults	46	1	Frequency and intensity	—
Pain Behavior Method (PBM)	Cognitively impaired (older) adults	5	1, 2, 4	Presence or absence	0–5
Discomfort Scale—Dementia of Alzheimer Type (DS-DAT)	Alzheimer patients	9	1, 2, 4, 6	4-point scale	0–27
DOLOPLUS2	Nonverbal or cognitively impaired older adults	10	1, 2, 3, 4, 5	4-point scale	0–30
Behavior Checklist	Cognitively impaired older adults	20	2, 3, 4, 5, 6	Presence or absence	—
Checklist of Nonverbal Pain Indicators (CNPI)	Cognitively impaired older adults	6	1, 2, 3, 4	Presence or absence	0–6
Assessment for Discomfort in Dementia (ADD)	Patients with moderate to severe dementia	5	1, 2, 3, 4, 5, 6	Presence or absence	—
Pain Assessment in Advanced Dementia (PAINAD)	Patients with (severe) dementia	5	1, 2, 3, 4, 6	3-point scale	0–10
Pain Assessment Tool in Confused Older Adults (PATCOA)	Confused older adults	9	1, 2, 4, 6	Presence or absence	0–9
Pain Assessment for the Dementing Elderly (PADE)	Patients with dementia	24	1, 2, 3, 4, 5, 6	4-point scale and multiple choice	24–96
Pain Assessment Checklist for Seniors with Limited Ability to Communicate (PACSLAC)	Seniors with a limited ability to communicate	60	1, 2, 3, 4, 5, 6	Presence or absence	0–60
Non-Communicative Patient's Pain Assessment Instrument (NOPPAIN)	Noncommunicative patients	17	1, 2, 3, 4, 5	6-point scale	0–30
Abbey Pain Scale	Patients with end-stage dementia	6	1, 2, 3, 4, 5, 6	4-point scale	0–18

Note. *1 = facial expression; 2 = motor behavior; 3 = social behavior or mood; 4 = vocalization; 5 = eat or sleep pattern; 6 = physiological indicators.

TABLE 2. Validity and Sensitivity

		Criterion Validity		Construct Validity		Sensitivity to Change
Instrument	No. of Studies	Concurrent: Correlation Between Two Pain Scales	Concurrent: Correlation Between Pain Scale and Proxy Pain Report	Convergent: Correlation Between Pain Scale and Self-report	Discriminant	Correlation Between Before and After Pain Medication
FACS	4	*n* = 58 PBM: *r* = .02 (guarding); .13 (bracing); and .41 (grimacing)	n = 26 AU (frequency) *r* = .62; AU (intensity) *r* = .73	*n* = 55; CAS: *r* = .05 *n* = 82; CAS/VDS *n* = 12; CAS	n = 26 Higher frequency and intensity of AUs during painful event *n* = 82 Significant higher scores during kneebending > standing > reclining	—
PBM	1	*n* = 58 FACS: *r* = .02–.41	—	*n* = 55; CAS: *r* = .11 (bracing); .21 (guarding); .30 (grimacing)	*n* = 82 Significant higher scores during kneebending > standing > reclining	—
DS-DAT	3	*n* = 19 PAINAD: *r* = .76	*n* = 19 *r* = .56 (pain by VAS) *r* = .81 (discomfort by VAS)	—	*n* = 82 Higher scores during fever episodes compared with baseline scores. *n* = 46 Higher scores during uncomfortable situations (care and transfer activities versus rest).	—
DOLOPLUS	2	—	—	*n* = 143 VAS: *r* = .65 *n* = 141 VAS/VRS/FPS: *r* = .31–.40	—	—
CNPI	1	—	—	*n* = 64 VDS: *r* = .30 at rest and *r* = .46 with movement	*n* = 26 Higher scores during movement compared to at rest	—
Behavior Checklist	1	—	—	—	—	*n* = 13 Lower scores after pain medication
ADD	2	—	—	—	—	*n* = 104 Fewer behavioral symptoms (*p* = .0001) *n* = 143 84% improved behavioral symptoms

TABLE 2. (continued)

Instrument	No. of Studies	Criterion Validity: Concurrent: Correlation Between Two Pain Scales	Criterion Validity: Concurrent: Correlation Between Pain Scale and Proxy Pain Report	Construct Validity: Convergent: Correlation Between Pain Scale and Self-report	Construct Validity: Discriminant	Sensitivity to Change: Correlation Between Before and After Pain Medication
PAINAD	1	n = 19; DS-DAT: r = .76	n = 19; r = .75 (pain by VAS); r = .76 (discomfort by VAS)	—	n = 19; Higher scores in unpleasant situations, $F(1,17)$ = 10.93	n = 19; Paired t test: lower scores after pain medication, $t(24)$ = 9.6, $p < .001$)
PATCOA	1	—	—	n = 116 VAS: r = .30	n = 116 Confusion: r = −.41	—
PADE	1	—	—	—	n = 40 Agitation (CMAI): r = .30–.42 Significantly higher scores in group "pain is clinical factor" No significant differences between with and without painful conditions.	—
PACSLAC	1	—	n = 40 r = .39 and r = .54	—	n = 40 Significant differences between pain, calm and stress situation; r = .80 between two pain situations.	—
NOPPAIN	1	—	—	—	Bradley–Terry model: deviance GFI = 18.14 (10)	—
Abbey Pain Scale	1	—	n = 61 Gamma = .59	—	—	n = 61; Significant difference before and after intervention ($p < .001$)

Note. FACS = Facial Activity Coding System; AU = Action Unit; CAS = Coloured Analogue Scale; VDS = Verbal Descriptor Scale; PBM = Pain Behavior Method; DS-DAT = Discomfort Scale—Dementia of Alzheimer Type; PAINAD = Pain Assessment in Advanced Dementia Scale; VAS = Visual Analogue Scale; DOLOPLUS = DOLOPLUS2; VRS = Verbal Rating Scale; FPS = Faces Pain Scale; CNPI = Checklist of Nonverbal Pain Indicators; ADD = Assessment of Discomfort in Dementia; PATCOA = Pain Assessment Tool in Confused Older Adults; PADE = Pain Assessment for the Dementing Elderly; CMAI = Cohen-Mansfield Agitation Inventory; PACSLAC = Pain Assessment Checklist for Seniors with Limited Ability to Communicate; NOPPAIN = Non-Communicative Patient's Pain Assessment Instrument.

properties of the PBM in 58 older adults with cognitive impairments during exercise at a rehabilitation facility. Overall interrater reliability was good but ranged from poor to good in terms of specific behaviors. No significant correlation was found between PBM scores and self-reported pain intensity. There were small to medium correlations in scores in the PBM and the FACS. The PBM seems easy to use in terms of conciseness and clear item definitions. However, the use of video recordings makes this scale less feasible in clinical practice. Also, only three items (guarding, grimacing, and bracing) were found to be reliable in an older adult population with cognitive impairments, and no information on sensitivity to change and cutoff scores are available.

The Discomfort Scale—Dementia of Alzheimer Type (DS-DAT; Hurley, Volicer, Hanrahan, Houde, & Volicer,

TABLE 3. Reliability, Feasibility, Clinical Utility

	No. of Studies	Interrater/Intrarater Reliability	Internal Consistency	Feasibility	Clinical Utility
FACS	2	Interrater: 43–93% agreement (frequency); Pearson's r = .82–.97 (intensity). Intrarater: 79–93% agreement (frequency); Pearson's r = .88–.97 (intensity)	—	Requires special training; completion is time consuming; extensive information on Web site	—
PBM	1	Interrater: ICC = .10–.87	—	Short scale; easy to use	—
DS-DAT	4	Interrater: Pearson's r = .61–.98; ICC = .74. Intrarater: ICC = .97	α = .77–.89	Intensive training with several practice moments needed; rater training is provided	—
DOLOPLUS2 Scale	1	Interrater: paired t test: no significant differences between physicians ($p < .001$)	α = .82	Easy to use; lexicon and instructions for use are available; takes a few minutes to complete	Scores >5 indicate presence of pain
CNPI	1	Interrater: 93% agreement; κ = .63–.82	α = .54 –.64	Easy to use	—
Behavior Checklist	1	—	—	—	—
ADD	1	Interrater: 76–100% agreement	—	Education sessions, didactic instructions and site visits needed before using the protocol	Assessment and treatment/ intervention plan in one
PAINAD	1	Interrater: Pearson's r = .82 and r = .97	α = .50–.72	Easy to use after minimal training; takes only a few minutes to complete	—
PATCOA	1	Interrater: 56–100% agreement; Spearman's rho = .16–1.00	α = .44	—	—
PADE	1	Interrater: ICC = .54–.96	α = .24–.88	Takes 5 to 10 minutes to complete	—
PACSLAC	1	Interrater: 94% agreement	α = .74–.92	Takes approximately 5 minutes to complete	—
NOPPAIN	1	Interrater: κ = .87; Pain level comparisons: 82–100% agreement	—	Mean of 8 minutes (range 3–15 minutes) to observe and complete; illustrations make use in practice easier	—
Abbey Pain Scale	1	Interrater: ICC = .63 and r = .44	α = .74–.81	Easy to use; takes 1 minute to complete	—

Note. FACS = Facial Activity Coding System; PBM = Pain Behavior Method; ICC = intra class correlation coefficient; DS-DAT = Discomfort Scale—Dementia of Alzheimer Type; CNPI = Checklist of Nonverbal Pain Indicators; ADD = Assessment of Discomfort in Dementia; PAINAD = Pain Assessment in Advanced Dementia Scale; PATCOA = Pain Assessment Tool in Confused Older Adults; PADE = Pain Assessment for the Dementing Elderly; PACSLAC = Pain Assessment Checklist for Seniors with Limited Ability to Communicate; NOPPAIN = Non-Communicative Patient's Pain Assessment Instrument.

1992) was developed to measure discomfort in patients with advanced dementia of the Alzheimer type. The scale consists of nine items (two positive and seven negative), among others: breathing, vocalization, facial expression, and body language. After 5 minutes of observation (minimum 30 minutes after an intervention), the observer records frequency, intensity, and duration of each item. Total score ranges from 0 (*no discomfort observed*) to 27 (*high level of discomfort observed*). Use of the DS-DAT requires in-depth training with several practice opportunities. Both an English language and a Dutch version have been validated in Alzheimer patients during presumably comfortable and uncomfortable situations or during acute illness, with sample size ranging from 19 to 97 (Hoogendoorn et al., 2001; Hurley et al., 1992; Miller et al., 1996; van der Steen, Ooms, van der Wal, & Ribbe, 2002). Interrater and intrarater reliability were good, test–retest correlation was moderate, and internal consistency was adequate to high. Construct validity showed a significant correlation between DS-DAT and self-report by question, but there was no significant correlation with a self-report thermometer. Furthermore, DS-DAT scores showed large correlations with the Pain Assessment in Advanced Dementia (PAINAD) observation scale, and DS-DAT showed higher scores in uncomfortable situations versus comfortable situations. In another study, construct validity was indicated by medium to large correlations with scores on the Pittsburgh Agitation Scale (Zieber, Hagen, Armstrong-Esther, &

Aho, 2005). Although different aspects of reliability and validity were evaluated in several studies, sensitivity to change and cutoff scores for discomfort need to be established. Also, DS-DAT is focused specifically on measuring discomfort instead of pain, which are not interchangeable because of different treatment protocols. The DS-DAT has a comprehensive scoring instruction and requires extensive training, which makes it less feasible.

The DOLOPLUS2 scale (Lefebvre-Chapiro & the DOLOPLUS group, 2001), based on the DOLOPLUS scale (Wary, Pandolfo, & Farnetti, 1993), was developed as a pain assessment instrument for nonverbal or cognitively impaired older adults. It is based on observations of behavior (somatic, psychomotor, and psychosocial) in 10 different situations that could reveal pain. Scoring in a multidisciplinary team is preferable. With item scores ranging from 0 to 3, the total score may range from 0 to 30 (Lefebvre-Chapiro & the DOLOPLUS group, 2001). The instrument seems easy to use and takes only a few minutes to complete. The DOLOPLUS group has conducted a validation study among 143 patients in geriatric or palliative care units. No diagnoses or specific painful conditions were described. Interrater reliability was good, test–retest reliability showed acceptable to good correlations, and internal consistency was also good. In terms of construct validity, a large correlation was found between the DOLOPLUS2 and self-report. Pautex et al. (2005) found medium correlations between the DOLOPLUS2 and different self-report measures in 120 patients with mild to severe dementia. Total scores of 5 or higher are suggested to indicate pain, which seems low for a 0–20 scale, and it remains unexplained how the cutoff score was reached. A comprehensive Web site (http://www.doloplus.com) and instructional videotapes on the DOLOPLUS2 scale are available, both in French and in English. However, the only validation work found was performed in a French-speaking and Norwegian-speaking population.

The Checklist of Nonverbal Pain Indicators (CNPI; Feldt, 2000) was designed to assess pain in cognitively impaired older adults in both acute and long-term-care settings. The checklist includes six behaviors (e.g., vocalizations, facial expression, body language). The total score is the number of the behaviors present, with total score ranging from 0 to 6. The authors suggest the instrument is easy to use in clinical practice. Feldt (2000) and Feldt, Ryden, and Miles (1998) studied the psychometric properties in 88 cognitively impaired and intact hospital patients with postoperative pain. Interrater reliability showed high agreement on the behaviors, the Kappa statistics were good to very good, and internal consistency was moderate. Medium correlations were found between the CNPI and self-report by the Verbal Descriptor Scale (VDS). Higher scores were found during activity versus at rest. Although different psychometric results are available, reliability and validity were not always as strong, and information on sensitivity to change and cutoff scores is not available.

The Behavior Checklist (Baker, Bowring, Brignell, & Kafford, 1996) was developed for cognitively impaired older adults and consists of 20 items (e.g., moaning, quiet, crying easily, rocking). Items are scored as present or absent. The authors conducted a double-blind intervention study in 13 cognitively impaired hospitalized patients showing pain (Baker et al., 1996). The Behavior Checklist showed improved comfort levels after acetaminophen administration, which points at good sensitivity to change. However, only one study about the Behavior Checklist was published, so more studies are required to confirm these preliminary results.

The Assessment of Discomfort in Dementia (ADD; Kovach, Weissman, Griffie, Matson, & Muchka, 1999) was designed to assess and treat discomfort and pain in people with moderate to severe dementia. The protocol consists of six behavioral categories to measure pain (facial expression, mood, body language, voice, behavior, and other). The observer scores the presence of behavioral symptoms. The authors studied the psychometric properties in 104 and 144 patients with dementia in long-term-care facilities (Kovach, Noonan, Griffie, Muchka, & Weissman, 2001; Kovach et al., 1999). Interrater reliability was good. Significant differences in behavioral symptoms, and increased use of scheduled analgesics and nonpharmacological comfort interventions point at sensitivity to change. In 88% of the cases, nurses reported the ADD protocol as somewhat helpful to very helpful. In contrast with most other observation scales, the ADD protocol combines observation of pain behavior with an ensuing treatment intervention plan for physical pain, affective discomfort, or both. More extensive validation studies are needed.

The Pain Assessment in Advanced Dementia Scale (PAINAD; Lane et al., 2003) was developed to measure pain in patients with severe dementia. The scale consists of five items (breathing, negative vocalizations, facial expression, body language, and consolability). Each item is rated 0, 1, or 2, resulting in a total score from 0 (*no pain*) to 10 (*maximal pain*). Rating is by severity of existing behavior, with 0 for normal behavior, 1 for more severe behavior, and 2 for most severe behavior. The instrument is easy to use after a 2-hour training session, and takes only a few minutes to complete. Nineteen patients with severe dementia were observed during rest, during a presumably pleasant activity (a visit), and during an unpleasant activity (e.g., a transfer; Warden, Hurley, & Volicer, 2003). Interrater reliability was good, and internal consistency ranged between inadequate and adequate. Large correlations were found between pain and discomfort measured by visual analogue scales and the PAINAD. Construct validity was confirmed by significant differences between the observed conditions and by large correlations with DS-DAT scores. PAINAD scores were significantly lower after analgesic administration. Overall, the sample size of the study was small, and somewhat low internal consistencies were found. Also, a cutoff score for pain was not provided.

The Pain Assessment Tool in Confused Older Adults (PATCOA; Decker & Perry, 2003) was developed and validated to assess postoperative pain in acutely confused older adults. It includes nine cues in four categories (vocalizations, behaviors, motor activities, and facial expressions). Total score is the number of items present. The psychometric properties were studied in 116 hospitalized cognitively intact older adults with postoperative pain. Interrater

reliability for the different items ranged from poor to good, and internal consistency was inadequate. As part of validity testing, the scale was correlated with a confusion scale and self-report, and medium correlations were shown. Overall, published psychometric properties were weak and limited, and a cutoff score is not provided. Also, the PATCOA has been validated only in cognitively intact older adults undergoing orthopedic surgery.

The Pain Assessment for the Dementing Elderly (PADE; Villanueva, Smith, Erickson, Lee, & Singer, 2003) was developed to assess pain in patients with dementia. The PADE is used to assess facial expressions, activities of daily living, and the caregiver's overall judgment of the resident's pain. The scale consists of 24 items in three categories: physical, global assessment, and functional. These items are rated on a visual analogue scale after 5 minutes of observation. In addition, 14 chart documentation data concerning the last 24 hours are recorded. The authors suggest that completing the PADE requires 5 to 10 minutes. Forty residents of long-term-care facilities with advanced levels of dementia who suffered potentially painful medical conditions were included in a validation study. Interrater reliability ranged from moderate to good, test–retest correlations ranged from low to good, and internal consistency coefficients for the subscales were inadequate (functional part) to adequate (physical part). Construct validity was indicated by medium correlations with scores on the Cohen-Mansfield Agitation Inventory. No significant differences between groups with and without painful conditions were found, but significantly higher scores were found in a group of patients in which pain was a significant clinical factor. A large range in reliability outcomes, dependent on the subscales, was shown in the PADE, and validity and sensitivity to change were tested insufficiently. Time to complete is not known, and the complexity of some items, the variety in scoring, and the need to review chart documentation of the last 24 hours make the PADE less feasible in clinical practice.

The Pain Assessment Checklist for Seniors with Limited Ability to Communicate (PACSLAC; Fuchs-Lacelle & Hadjistavropoulos, 2004) consists of 60 items in four categories (facial expressions, activity/body movements, social and personality changes, and other). The category *other* includes a variety of pain behaviors (e.g., appetite or sleeping changes). Subscale scores are derived from counting the checkmarks (present or absent) in each column. Summation of the four subscale totals generates a total pain score ranging from 0 to 60. In a preliminary validation study, patients with cognitive impairments were observed during two painful events, a nonpainful distress event, and at rest (Fuchs-Lacelle & Hadjistavropoulos, 2004). Interrater reliability was very good, and internal consistency was adequate to high. Concurrent validity was assessed by correlating nurses' ratings with ratings for two painful events, which resulted in medium correlations. Furthermore, significant differences were found between painful and pain-free situations. Mean completion time was 5 minutes, and clear instructions for use are available. Preliminary validation results show reasonable properties, but more validation studies are needed, in particular, prospective ones, and cutoff scores need to be tested.

The Non-Communicative Patient's Pain Assessment Instrument (NOPPAIN; Snow et al., 2004) focuses on pain assessment by caregivers in patients with dementia. The scale rates the following aspects: pain in response to activities of daily living, vocalization and facial expression, bracing and restlessness, and a global rating of pain for that day. Patients are observed during care activity, and the caregiver completes the NOPPAIN afterwards. Pictures facilitate use for caregivers with poor English language skills. Validity was studied by Snow et al. (2004) using video recordings of an actress portraying a bed-bound patient with severe dementia during caregiving. Interrater reliability was good. Construct validity comparing ratings of caregivers with six videotaped pain levels was moderate. Especially differentiating between mild and moderate pain levels was difficult for the caregivers. A mean of 8 minutes (range 3–15) of caring and observing was found, and less than 30 seconds of completion time. Overall, sensitivity to change and cutoff scores were not tested. Also, use of an actress portraying a patient with dementia as the gold standard of pain has not been described in the literature as a reliable method.

The Abbey Pain Scale (Abbey et al., 2004) was developed to measure pain in people with end-stage dementia. Six behavioral indicators are included (vocalization, facial expression, change in body language, behavioral change, physiological change, and physical change). While observing a patient, the observer rates the indicators as absent, mild, moderate, or severe (0–3). The total pain score thus ranges from 0 to 18. Total scores 0–2 indicate no pain, 3–7 indicate mild pain, 8–13 indicate moderate pain, and 14 or higher indicate severe pain. Type of pain (chronic, acute, or acute on chronic) can be noted. The Abbey Pain Scale is easy to complete and takes less than 1 minute. Abbey et al. (2004) validated the scale for 61 patients with end- or late-stage dementia living in residential aged care facilities. The Abbey Pain Scale was completed when staff judged patients to be in pain. Interrater reliability was modest, and reliability analysis provided adequate to high internal consistencies. Concurrent validity showed a significant correlation between the pain score and a holistic pain assessment score by nurses. Furthermore, significant differences were found between pain scores before and after administration of analgesics. Overall, score ranges indicative of level of pain are given, but it is not clear how these were reached. Also, clear scoring instructions and item definitions are lacking.

Discussion

A great variety of observation scales have been developed to measure varying types of pain in varying groups of patients. Also, psychometric evaluation is not uniform among studies. For these reasons, it is difficult to determine the optimal assessment.

The extensive validation studies of the FACS, the PACSLAC, the DS-DAT, and the PAINAD show the most promising outcomes. Unfortunately, cutoff scores for these scales have not been established. Cutoff scores are important in deciding whether interventions to alleviate pain are required. They also appear to motivate nurses to assess

their patients' pain. In view of the extensive training and analyses required for the FACS, the complexity of scoring and interpretation of the DS-DAT, and the large number of items in the PACSLAC, the PAINAD seems the best feasible scale for clinical practice.

Effective scale development in persons who are not able to report pain themselves is a challenge. Because no gold standard is available to validate an instrument in this population, one has resorted to *silver standards*—proxy reports from caregivers or family members. Although it may be effective to use judgments of caregivers and family members who are familiar with the patient (Mentes, Teer, & Cadogan, 2004), low agreement between patient's self-report and proxy reports was found, especially in the more severely cognitively impaired (Werner, Cohen-Mansfield, Watson, & Pasis, 1998). Caregivers and family members typically tend to underestimate chronic pain. Using self-report as a validation method requires a subsample of mildly cognitively impaired or intact persons because most moderately to severely cognitively impaired people are not able to report pain reliably.

Overall, varying results were found between self-report and pain observation (Hadjistavropoulos et al., 2000; Lefebvre-Chapiro & the DOLOPLUS group, 2001; McCahon, Strong, Sharry, & Cramond, 2005). Additionally, results concerning proxy reports and self-reports should be interpreted carefully, and other standards should be considered (e.g., comparisons with existing observation scales to test concurrent validity). Thus, in developing new scales, observing patients, in particular during care activities when pain behavior is most likely or those patients who are likely to be in pain (e.g., postoperative patients or osteoarthritic patients), could be useful as part of the validation process. Furthermore, to achieve sufficient sensitivity to change, it is essential to demonstrate differences in scores between painful and pain-free situations, and before and after changes in prescribed analgesics or other pain-relieving therapies. As nurses have limited time, a practical scale should be concise and easy to use. Clear definitions about type of pain and a scoring manual would make the assessment more useful. Information on the time needed to complete training sessions should be available.

Future Goals

Differences in pain behavior or pain experience between patients with different types of dementia or different levels of cognitive functioning have been reported (Benedetti et al., 1999; Scherder, Sergeant, & Swaab, 2003). Therefore, results obtained in overall dementia groups must be interpreted carefully. Future research focusing on pain thresholds and pain tolerance levels in the older adults could give valuable information for the validation of a new scale. If the underlying pain thresholds and tolerances of different diseases are known, these levels can be used to measure pain more efficiently. The pain scores are then more reliable and with those, a validation study is more reliable. Ideally, a new pain scale should be tested in various settings to improve external validity also.

Lastly, in order to achieve less pain and better quality of life, pain scales should be linked with a treatment algorithm based on well-calculated cutoff scores. Importantly, pain should be assessed also after treatment to determine efficacy of treatment.

Conclusions

A reliable, valid, and feasible pain scale enables to treat pain more adequately. As different methods of evaluating reliability and validity are used in the available scales, and different aims (e.g., type pain) are pursued, they cannot be compared easily. More specific research is needed to develop a linked algorithm for pain treatment based on cutoff scores. Furthermore, not only nurses but also physicians and other disciplines working with older adults need to gain more knowledge about identifying chronic pain behavior and treatment of pain. Pain observation scales should at least be tailored to the unique characteristics and needs of the older adult with communication difficulties. Until a reliable and valid observation scale has been developed, patients with one or more probable painful diagnoses who are not able to communicate should be treated as if they are in pain. In addition, the most promising pain scales should be used to observe the presence of specific pain behaviors. ▼

Accepted for publication August 14, 2006.

Thank you to J. Hagoort for help in preparing the manuscript.

Corresponding author: Rhodee van Herk, MA, Pain Expertise Center, Erasmus MC, V019, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands (e-mail: r.vanherk@erasmusmc.nl).

References

Abbey, J., Piller, N., De Bellis, A., Esterman, A., Parker, D., Giles, L., et al. (2004). The Abbey pain scale: A 1-minute numerical indicator for people with end-stage dementia. *International Journal of Palliative Nursing*, *10*(1), 6–13.

American Geriatrics Society Panel on Persistent Pain in Older Persons. (2002). The management of persistent pain in older persons. *Journal of the American Geriatrics Society, 50*(6 Suppl.), S205–S224.

Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.

Baker, A., Bowring, L., Brignell, A., & Kafford, D. (1996). Chronic pain management in cognitively impaired patients: A preliminary research project. *Perspectives, 20*(2), 4–8.

Benedetti, F., Vighetti, S., Ricco, C., Lagna, E., Bergamasco, B., Pinessi, L., et al. (1999). Pain threshold and tolerance in Alzheimer's disease. *Pain*, *80*(1–2), 377–382.

Closs, S. J., Barr, B., Briggs, M., Cash, K., & Seers, K. (2004). A comparison of five pain assessment scales for nursing home residents with varying degrees of cognitive impairment. *Journal of Pain and Symptom Management*, *27*(3), 196–205.

Closs, S. J., Cash, K., Barr, B., & Briggs, M. (2005). Cues for the identification of pain in nursing home residents. *International Journal of Nursing Studies*, *42*(1), 3–12.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale: Lawrence Erlbaum.

Craig, K. D., Prkachin, K. M., & Grunau, R. V. E. (1992). The facial expression of pain. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 257–276). New York: Guilford Press.

Decker, S. A., & Perry, A. G. (2003). The development and testing of the PATCOA to assess pain in confused older adults. *Pain Management Nursing*, *4*(2), 77–86.

Ekman, P., & Friesen, W. (1978). *Investigator's guide to the Facial Action Coding System*. Palo Alto: Consulting Psychologists Press.

Feldt, K. S. (2000). The checklist of nonverbal pain indicators (CNPI). *Pain Management Nursing*, *1*(1), 13–21.

Feldt, K. S., Ryden, M. B., & Miles, S. (1998). Treatment of pain in cognitively impaired compared with cognitively intact older patients with hip-fracture. *Journal of the American Geriatric Society*, *46*(9), 1079–1085.

Ferrell, B. A., Ferrell, B. R., & Rivera, L. (1995). Pain in cognitively impaired nursing home patients. *Journal of Pain and Symptom Management*, *10*(8), 591–598.

Fuchs-Lacelle, S., & Hadjistavropoulos, T. (2004). Development and preliminary validation of the Pain Assessment Checklist for Seniors with Limited Ability to Communicate (PACSLAC). *Pain Management Nursing*, *5*(1), 37–49.

Hadjistavropoulos, H. D., Craig, K. D., Martin, N., Hadjistavropoulos, H. D., & McMurtry, B. (1997). Toward a research outcome measure of pain in frail elderly in chronic care. *Pain Clinic*, *10*(2), 71–79.

Hadjistavropoulos, T., LaChapelle, D. L., Hadjistavropoulos, H. D., Green, S., & Asmundson, G. J. (2002). Using facial expressions to assess musculoskeletal pain in older persons. *European Journal of Pain*, *6*(3), 179–187.

Hadjistavropoulos, T., LaChapelle, D. L., MacLeod, F. K., Snider, B., & Craig, K. D. (2000). Measuring movement-exacerbated pain in cognitively impaired frail elders. *Clinical Journal of Pain*, *16*(1), 54–63.

Hoogendoorn, L. I., Kamp, S., Mahomed, C. A., Ader, H. J., Ooms, M. E., & van der Steen, J. T. (2001). [The role of observer for the reliability of Dutch version of the Discomfort Scale—Dementia of Alzheimer Type (DS-DAT)]. *Tijdschrift voor Gerontologie en Geriatrie*, *32*(3), 117–121.

Hurley, A. C., Volicer, B. J., Hanrahan, P. A., Houde, S., & Volicer, L. (1992). Assessment of discomfort in advanced Alzheimer patients. *Research in Nursing & Health*, *15*(5), 369–377.

Jensen, M. P. (2003). Questionnaire validation: A brief guide for readers of the research literature. *Clinical Journal of Pain*, *19*(6), 345–352.

Keefe, F., & Block, A. (1982). Development of an observation method for assessing pain behavior in chronic low back pain patients. *Behavior Therapy*, *13*, 363–375.

Kovach, C. R., Noonan, P. E., Griffie, J., Muchka, S., & Weissman, D. E. (2001). Use of the assessment of discomfort in dementia protocol. *Applied Nursing Research*, *14*(4), 193–200.

Kovach, C. R., Weissman, D. E., Griffie, J., Matson, S., & Muchka, S. (1999). Assessment and treatment of discomfort for people with late-stage dementia. *Journal of Pain and Symptom Management*, *18*(6), 412–419.

LaChapelle, D. L., Hadjistavropoulos, T., & Craig, K. D. (1999). Pain measurement in persons with intellectual disabilities. *Clinical Journal of Pain*, *15*(1), 13–23.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lane, P., Kuntupis, M., MacDonald, S., McCarthy, P., Panke, J. A., Warden, V., et al. (2003). A pain assessment tool for people with advanced Alzheimer's and other progressive dementias. *Home Healthcare Nurse*, *21*(1), 32–37.

Lefebvre-Chapiro, S., & the DOLOPLUS group. (2001). The DOLOPLUS2 scale—Evaluating pain in the elderly. *European Journal of Palliative Care*, *8*, 191–194.

Manfredi, P. L., Breuer, B., Meier, D. E., & Libow, L. (2003). Pain assessment in elderly patients with severe dementia. *Journal of Pain and Symptom Management*, *25*(1), 48–52.

McCahon, S., Strong, J., Sharry, R., & Cramond, T. (2005). Self-report and pain behavior among patients with chronic pain. *Clinical Journal of Pain*, *21*(3), 223–231.

Mentes, J. C., Teer, J., & Cadogan, M. P. (2004). The pain experience of cognitively impaired nursing home residents: Perceptions of family members and certified nursing assistants. *Pain Management Nursing*, *5*(3), 118–125.

Miller, J., Neelon, V., Dalton, J., Ng'andu, N., Bailey, D., Jr., Layman, E., et al. (1996). The assessment of discomfort in elderly confused patients: A preliminary study. *Journal of Neuroscience Nursing*, *28*(3), 175–182.

Morrison, R. S., & Siu, A. L. (2000). A comparison of pain and its treatment in advanced dementia and cognitively intact patients with hip fracture. *Journal of Pain and Symptom Management*, *19*(4), 240–248.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psyhometric theory*. New York: McGraw-Hill.

Pautex, S., Herrmann, F., Le Lous, P., Fabjan, M., Michel, J. P., & Gold, G. (2005). Feasibility and reliability of four pain self-assessment scales and correlation with an observational rating scale in hospitalized elderly demented patients. *Journal of Gerontology. Series A, Biological Sciences and Medical Sciences*, *60*(4), 524–529.

Polit, D. F., & Hungler, B. P. (1987). *Nursing research: Principles and methods*. Philadelphia: Lippincott.

Scherder, E. J., Sergeant, J. A., & Swaab, D. F. (2003). Pain processing in dementia and its relation to neuropathology. *Lancet Neurology*, *2*(11), 677–686.

Snow, A. L., Weber, J. B., O'Malley, K. J., Cody, M., Beck, C., Bruera, E., et al. (2004). NOPPAIN: A nursing assistant-administered pain assessment instrument for use in dementia. *Dementia and Geriatric Cognitive Disorders*, *17*(3), 240–246.

van der Steen, J. T., Ooms, M. E., van der Wal, G., & Ribbe, M. W. (2002). [Measuring discomfort in patients with dementia. Validity of a Dutch version of the Discomfort Scale—dementia of Alzheimer type (DS-DAT)]. *Tijdschrift voor Gerontologie en Geriatrie*, *33*(6), 257–263.

Villanueva, M. R., Smith, T. L., Erickson, J. S., Lee, A. C., & Singer, C. M. (2003). Pain Assessment for the Dementing Elderly (PADE): Reliability and validity of a new measure. *Journal of the American Medical Directors Association*, *4*(1), 1–8.

Warden, V., Hurley, A. C., & Volicer, L. (2003). Development and psychometric evaluation of the Pain Assessment in Advanced Dementia (PAINAD) scale. *Journal of the American Medical Directors Association*, *4*(1), 9–15.

Wary, B., Pandolfo, J., & Farnetti, S. (1993). Ce vieillard a t-il mal? In D. J. Roy & C.-H. Rapin (Eds.), *Douleur et antalgie. Annales de soins palliatifs* (pp. 89–98). Montreal: Institut de recherches cliniques de Montréal, Centre de bioéthique.

Werner, P., Cohen-Mansfield, J., Watson, V., & Pasis, S. (1998). Pain in participants of adult day care centers: Assessment by different raters. *Journal of Pain and Symptom Management*, *15*(1), 8–17.

World Health Organization. (2002). *Active ageing: A policy framework*. Geneva.

Zieber, C. G., Hagen, B., Armstrong-Esther, C., & Aho, M. (2005). Pain and agitation in long-term care residents with dementia: Use of the Pittsburgh Agitation Scale. *International Journal of Palliative Nursing*, *11*(2), 71–78.