Comparing the Psychometric Properties of the Checklist of Nonverbal Pain Behaviors (CNPI) and the Pain Assessment in Advanced Dementia (PAIN-AD) Instruments

Mary Ersek, PhD, RN, FAAN[Associate Professor],
University of Pennsylvania School of Nursing, Philadelphia, PA 19104-6096

Keela Herr, RN, PhD, FAAN, AGSF[Professor & Chair],
Adult & Gerontology Nursing, University of Iowa College of Nursing, Iowa City, IA

Moni Blazej Neradilek, MS[Statistical Analyst],
The Mountain-Whisper-Light Statistical Consulting, Seattle, WA

Harleah G. Buck, PhD, RN, CHPN[Post-Doctoral Research Fellow], and
Hartford Center of Geriatric Nursing Excellence, NewCourtland Center for Transitions and Health, University of Pennsylvania School of Nursing, Philadelphia, PA

Brianne Black, RN, BSN[John A Hartford Pre-doctoral Scholar]
University of Iowa College of Nursing, Iowa City, IA

Abstract

Objective—To examine and compare the psychometric properties of two common observational pain assessment tools used in persons with dementia.

Design—In a cross-sectional descriptive study nursing home (NH) residents were videotaped at rest and during a structured movement procedure. Following one training session and one practice session, two trained graduate nursing research assistants independently scored the tapes using the two pain observation tools.

Setting—Fourteen nursing homes in Western Washington State participating in a randomized controlled trial of an intervention to enhance pain assessment and management.

Participants—Sixty participants with moderate to severe pain were identified by nursing staff or chosen based on the pain items from the most recent Minimum Data Set assessment.

Measures—Checklist of Nonverbal Pain Indicators (CNPI) and the Pain Assessment in Advanced Dementia (PAINAD), demographic and pain-related data (Minimum Data Set), nursing assistant reports of participants' usual pain intensity, Pittsburgh Agitation Scale (PAS).

Results—Internal consistency for both tools was good except for the CNPI at rest for one rater. Inter-rater reliability for pain presence was fair (K = 0.25 for CNPI with movement; K = 0.31 for PAINAD at rest) to moderate (K = 0.43 for CNPI at rest; K = 0.54 for PAINAD with movement). There were significant differences in mean CNPI and PAINAD scores at rest and during movement, providing support for construct validity. However, both tools demonstrated marked floor effects, particularly when participants were at rest,

Conclusions—Despite earlier studies supporting the reliability and validity of the CNPI and the PAINAD, findings from the current study indicate that these measures warrant further study with

Corresponding author: Mary Ersek, PhD, RN, FAAN, Associate Professor, University of Pennsylvania, School of Nursing, Room 329, 418 Curie Bld., Philadelphia, PA 19104-6096, Office phone: (215) 746-3563, Office fax:(215) 573-7507, ersekm@nursing.upenn.edu.

clinical users, should be used cautiously both in research and clinical settings and only as part of a comprehensive approach to pain assessment.

Introduction

An individual's report of his or her pain generally is recognized as the cornerstone of pain measurement and assessment.(1–4) In patients with advanced dementia, however, this standard often cannot be met, as cognitive and self-expressive abilities decline. When pain self-report is unavailable, other indications of pain must be sought. One commonly used strategy is to identify, measure and document behaviors that are associated with pain.

Several observational pain measures for persons with advanced dementia have been developed, and recent reviews indicate no one tool can be recommended broadly for use across care settings and populations. Authors also noted the need for further evaluation and study of the tools.(5–9)

One of the primary differences among the tools is whether or not the person needs to be observed over time by a caregiver who is familiar with the person or an assessment is made based on a shorter period (for example, during the care provider's shift). The advantage of the former approach is the ability to detect subtle changes in behavior such as appetite or sleep patterns that may signal the presence of an increase in pain. Tools that are comprised of fewer, but more obvious pain behaviors that are observed over a briefer period, may be more appropriate for more frequent pain assessment, monitoring of responses to therapy, or for use in acute care or clinic settings(6) These tools may also be used in conjunction with more comprehensive tools, which may serve as screening measures.(6)

Of the available behavioral observation tools, the Checklist of Nonverbal Pain Indicators (CNPI) and the Pain Assessment in Advanced Dementia (PAINAD) are notable in that they are brief, require limited time to administer, and do not rely on extensive knowledge of the patient over time. They both have undergone psychometric testing and are used clinically and in research.(10–24) These tools have particular value in that caregivers, including nursing assistants and other unlicensed personnel can be trained to use them.(14) A recent review of pain observation tools supported further research on the CNPI and another review recommended the PAINAD for clinical use.(25,26)

Despite the potential utility of these tools for clinical practice and research, additional investigation is needed. Clinicians are faced with decisions related to which tool is best for their setting and patient population. In the nursing home, in particular, the CNPI and PAINAD have been tested but no comparison of tool performance using the same procedures and sample has been conducted.

The purpose of this study was to examine and compare the reliability and validity of the CNPI and PAINAD in detecting pain in older persons with dementia and known persistent pain. Although the authors had no *a priori* hypotheses about the superiority of one tool over the other, the analyses were guided by the assumptions regarding strong psychometric properties for measurement tools including: 1) Internal consistency measured by Cronbach's alpha coefficient will be 0.70 or higher for both scales; 2) inter-rater agreement between two trained research assistants will be moderate or better for both scales; 3) scores will be significantly higher on both instruments when measuring pain-related behaviors during

movement versus at rest; 4) there will be positive associations between scores on a measure of agitation with scores on each pain tool.

Methods

Data for this study were collected as part of an ongoing randomized controlled trial evaluating the effectiveness of a pain management algorithm and intensive diffusion strategies in nursing homes. All study procedures were approved by the Swedish Medical Center Institutional Review Board and the participants' designated proxies provided written consent.

Participants

Participants were residents of 14 Western Washington nursing homes who were eligible to participate in the parent study if they were 65 years or older and had experienced moderate to severe pain during the week prior to baseline data, as determined by self- or surrogate report of pain. In addition to having moderate to severe pain, residents were eligible if they were receiving long-term care at the facility and were expected to live at least 6 months (the length of the study). Residents who were receiving hospice services during the recruitment period were excluded because of their prognosis; in addition, their pain management would be directed by the hospice team rather than the nursing home staff..

Identification of residents with pain was achieved using the following three sequential steps: 1) Every unit manager (or other licensed nurse with care planning responsibilities) reviewed a roster of residents on the unit and identified those she or he assessed as meeting study criteria; 2) A list of all residents with MDS pain scores (i.e., Section J2b. Pain intensity) of "2 = Moderate Pain" or "3 = Times when pain is horrible or excruciating" was obtained; 3) Medical records of all residents who were not identified as having pain were reviewed and evaluated for indications of pain through progress or other clinical notes that mentioned the resident had moderate to severe pain and/or presence of analgesic orders.

Residents who were identified as having pain through the first three steps were interviewed by trained research staff to verify they had experienced moderate to severe pain in the previous seven days using the Iowa Pain Thermometer (IPT). In cases where residents were unable to self-report pain, the certified nursing assistant (CNA) working with the resident was queried as to whether or not the resident exhibited signs of pain. Using these methods participants were entered into the study regardless of their cognitive status or ability to self-report.

Only those participants who were nonverbal or unable to provide reliable self-report of pain were videotaped and included in the current study. Reliability was assessed in one of two ways. When the study began, participants were handed the IPT and asked to imagine the worst pain they had ever experienced. A participant was evaluated as providing a reliable response if he or she could report a time when they experienced their worst pain and pointed to the top third of the IPT (all reported having experienced severe pain as their worst). Some participants, however, struggled with this method even though they otherwise seemed capable of responding to other questions in the interview. Through discussions among the data collection staff and investigators, an alternate approach was taken. For the alternate method, the research assistant who interviewed the participant reviewed the participants' answers for *usual, worst, least* and *current pain*. If all responses were consistent (i.e., worst pain was greater than *least* pain) then the participant was assessed as providing reliable answers. If any response was not consistent, the research assistant queried the participant further for clarification; for example, the participant might be asked how his *least* pain was higher than his *current* pain. If the participant changed his response to be logically

consistent or if he stated that at present he was experiencing *least* pain, then the response was marked "reliable." If the participant was not able to identify the discrepancy, then the response was marked "unreliable."

Study Procedures

Video-taping procedure—As part of the ongoing clinical trial, participants who were unable to provide reliable self-report of pain were videotaped at rest and during activity, according to a standardized protocol. Research staff identified a CNA who had recently cared for and was familiar with the participant to participate in the filming. All videotaping was conducted in the participant's room. The participant was first filmed for approximately one minute while sitting or lying in bed, in other words, at rest. Following this period, the CNA was guided through a series of basic care activities: transferring the participant from the bed to a chair (or vice versa), removing or putting on an article of clothing (e.g., a sweater or jacket) to the upper body, and removing or putting on the participant's sock and shoes. Depending on the participant's abilities, they performed these activities with verbal prompting only, with some assistance, or total assistance from the CNA. When necessary, adaptive equipment such as walkers or lifts were used to facilitate safe transfer. The decision to videotape participants at rest before movement was based on evidence that movement could likely increase pain which may not resolve during a rest period that followed immediately.

Two video cameras were used. Camera 1 was positioned to frame a close-up of the participant's face to maximize observation of facial expression. Camera 2 was positioned to capture the entire body looking for behaviors such as bracing, guarding, and rubbing. Raters viewed the videotapes from both camera angles and scored each item on the PAINAD and CNPI by marking whether the behaviors were seen on either camera angle. The average videotape length, which included both camera angles, was 11.5 minutes (range 6—19.5 minutes).

Videotape scoring—Two training sessions were conducted and included two nurse researchers (ME, KH) and two graduate nursing students. The first session was 1 ½ hours and included review of each tool and its components and scoring procedures. All four individuals reviewed a sample of four videos to score, compared and discussed any differences in behavior interpretation. A second session lasting 1 hour was held one week after each RA rated four more videos to compare ratings and identify any discrepancies in scoring approaches.

Two trained RAs viewed the videos of each participant from both close up and wide angles. All videos were scored with one tool first, and then viewed a second time to score with the second tool. There were at least two weeks between the first and second viewing. In an effort to minimize biases, half of the videos were scored using the CNPI first, followed by the PAINAD and the other half were scored in reverse order. The videos were compiled on CDs in order of filming by site. The videos were viewed in a random order. Training manuals were not available for the CNPI or the PAINAD; however two authors (ME, KH) who were familiar with the tools trained the RAs in the tool's use. For both instruments, the RAs were instructed to only score behaviors specified on the tool. Each patient was scored at rest and with movement. For the CNPI, behaviors were scored as either present or not present. On the PAINAD, behaviors are given a numerical score based on the severity of the behavior.

For this study, a facial grimace included the majority of the face and the furrowing of the eyebrows while a frown included a sad appearance. Noises were considered anything that could not be clarified into words and words were actual pain related statements or remarks.

Bracing was described as a clear sign of gripping the side rail, staff or other object and could not be confused with rubbing an area, showing what area hurt or unintentional gripping. In general, the RAs were instructed to refrain from drawing inferences about the behaviors and assume that present behaviors were related to pain.

Measures

Completion of all study instruments for each participant occurred within 3 days. The videotaping also occurred in this time frame. Demographic data were gathered using the admission Minimum Data Set (MDS v. 2.0), a federally-mandated assessment and care planning tool.(27) In a few instances, data in the MDS were incomplete and research staff examined other parts of the medical record (e.g., social work assessment) to locate the data.

Cognitive Performance Scale (CPS)—The CPS was used to characterize the level of cognitive impairment of the sample. This 5-item instrument, derived from the most recent MDS, measures short and long-term memory, cognitive skills for decision-making, communication, and independence in eating. Following standard decision rules, raters assign a summary score of 0 (intact) to 6 (very severe impairment). Morris et al reported that scores on the CPS were significantly associated with the Mini-Mental State Examination (MMSE), nursing assessments, and neurological diagnoses of Alzheimer's disease and other dementias.(28) Overall inter-rater reliability of all items is acceptable at 0.85.(28)

Proxy pain reports from CNAs—Proxy pain reports for participating participants were obtained from CNAs using the Iowa Pain Thermometer (IPT). The IPT uses a graphic representation of a thermometer in which the base is white and becomes increasingly red as one moves up. The base is anchored with the words "no pain," and the top of the thermometer is anchored with "the most intense pain imaginable." Thirteen evenly spaced circles, corresponding with numeric values from "0" to "12," are placed between the thermometer the verbal descriptors. The CNA who was caring for the resident was asked to "rate the resident's pain during the past week." The trained data collectors attempted to elicit proxy reports from CNAs who were familiar with the resident. Data collectors asked CNAs how many times they had worked with the resident in the past week. If the CNA had not cared for the resident during that period, then the research staff sought input from a CNA who had recently cared for the resident. Several studies have shown that the IPT is reliable and valid in older adults.(29–32) This tool was used for the CNA proxy measure so as to be consistent with the one used by self-reporting study participants.(33)

Pain-related diagnoses—Painful diagnoses were collected from the medical record using the MDS and the participant's medical problem list. Two physicians and one gerontological nurse practitioner, all with expertise in geriatrics and pain management, reviewed a list of diagnoses from the two sources and judged whether or not the condition consistently (i.e., in at least 75% of participants with the condition) was associated with pain. To be included as a painful diagnosis, at least two of the three experts had to agree. A final Painful Diagnoses score was calculated for each participant by adding all diagnoses evaluated as painful by the expert group. Diagnoses that appeared both in the MDS and on the problem list were counted only once.

Agitation—Agitation has been associated with pain in persons with cognitive impairment. (34–36) As such, agitation measures can provide evidence of construct validity for observational pain tools. Agitation was measured using the Pittsburgh Agitation Scale (PAS), which is an observer-rating of four groups of behaviors: aberrant vocalization, motor agitation, aggressiveness and resistance to care. Respondents are asked to rate the highest intensity score for each behavior group that was observed during the past week. The four

items are scored on a "0" to "4" scale, with "0" indicating the behavior was not present and "4" indicating the most intense demonstration of the behavior. Overall scores range from 0—16, with higher scores denoting high levels of agitation. The PAS has acceptable reliability and validity (37) and has been found to be significantly associated with behavioral and surrogate pain measures in nursing home residents with moderate to severe cognitive impairment.(35) For the current study, the PAS was completed by the registered nurse responsible for overseeing the participant's overall plan of care. The timeframe for the responses on the PAS was "in the past week."

Checklist of Nonverbal Pain Indicators (CNPI)—The CNPI was adapted from the University of Alabama-Birmingham Pain Behavior Scale (UAB-PBS).(38) It has six clusters of behaviors (vocalizations, facial grimaces, bracing, rubbing restlessness and verbal complaints) which are scored as 'present' or 'absent' under two conditions: 'rest' and 'with movement.' Summing the number of checked items in each condition yields subscale scores (range 0–6) and a score (range 0–12).(11) The CNPI initially was evaluated in hospitalized elders and demonstrated adequate reliability and validity.(11) Later studies provide support for the tool's construct validity,.(13,14) test-retest reliability,(14) and inter-rater reliability,(14) in nursing home residents with persistent pain. In the initial testing, Feldt (11) did not offer guidance in the interpretation of summed scores, although two subsequent studies reported significant, moderate positive correlations between the number of pain indicators and self-reported pain intensity.(13,14)

Pain Assessment in Advanced Dementia (PAINAD)—The PAINAD is derived from two different pain observation scales: the Discomfort Scale—Dementia of Alzheimer Type (DS-DAT) (39) and the FLACC Scale for scoring postoperative pain in young children.(40) and aims to provide an easy-to-use tool that allows for the quantification of pain severity on a 0—10 scale. (15) Five items (breathing, vocalization, facial expression, body language, consolability) are each scored on a 0—2 scale, with higher scores indicating greater pain intensity. Follow up studies provides further support for the tool's validity and reliability. (16,20–22) The tool has been evaluated in both nursing home and acute care settings.(17,24) The PAINAD has good internal consistency with lowest associations for the indicator of breathing.(16,21) Strong inter-rater reliability and test-retest reliability has been reported (7,16,18,21) The developers do not provide guidance in interpreting the score of the PAINAD, but studies suggest that increasing numbers on the tool indicate higher levels of pain severity.

Statistical Analysis

Patients' demographics and pain variables were presented as means and standard deviations (SDs) for ordinal and categorical variables or as percentages for categorical variables. Floor effects were assessed by dichotomizing items and total scores at rest and with movement into "No Pain" indicators (score = 0) or "Pain Present" (any score > 0). Internal consistency was assessed by the Cronbach's alpha coefficients for both scales. The inter-rater reliability of the two research assistants was assessed. Cohen's kappa was used to quantify reliability for pain presence and the intra-class correlation coefficient (ICC) was used to quantify reliability of the total score. Cohen's kappa for pain presence represented the chance-adjusted likelihood of between-rater agreement in the indicator of pain presence. ICC represented the relative comparison of the between subject variation and between rater variation (i.e. ability to distinguish between patients in the presence of measurement error due to different interpretation by different raters). The 95% confidence intervals for Cohen's kappa and ICC were calculated using the adjusted bootstrap 95th percentile interval.(41)

Construct validity is the extent to which an instrument measures the specific concept or phenomenon of interest. Construct validity is evaluated in several ways including the examination of: 1) the correlation of the scale with measures of the same construct using scales that are already recognized as "gold standard" measures (criterion validity), 2) associations between the scale and other instruments measuring similar (convergent validity) and different (discriminant validity) concepts, and 3) correlations between the measure across samples or conditions (e.g., low and high pain states).(42). Discriminant validity for each instrument also was evaluated by comparing scores during rest and movement using the paired t-test.. We assessed convergent validity by computing the Spearman correlations between each of the two tools and the Pittsburgh Agitation Scale (PAS); we hypothesized that each would be significantly associated with PAS, given previous research that demonstrated significant associations between pain and agitation in older adults with advanced dementia.(34,35,43)

Results

Sample

Participants were predominately white, nonHispanic (93%) females (88%) with a mean age of 89 (SD=6.8) years. The mean CPS was 3.9 (SD=1.2) indicating moderate to moderately-severe cognitive impairment (Table 1).

Eighty percent had at least one pain-related diagnosis documented on the MDS or problem list, with almost 62% having a documented diagnosis of osteoarthritis. Only 38% were identified on the MDS as having pain on a daily or less than daily basis, with 48% (n=11) of those with pain judged to experience moderate to excruciating pain. CNAs reported a mean "usual pain" score of 3.08 (SD=2.6) for these participants (Table 1).

CNPI

Tables 2 and 4 summarize the CNPI values among the participants. Mean total subscale CNPI scores were 0.9 (SD=2.0) and 0.8 (SD=1.7) for rest and 1.9 (SD=2.2) and 2.0 (SD=1.6) for movement, depending on the rater. The distribution of the scores was right-skewed, with many scores clustered around zero. For the resting condition, 80—92% of total subscales scores were zero and with movement, 15—42% of subscale scores were zero.

Cronbach's alpha coefficients for total CNPI at rest were .97 and .92, and .74 and .90 with movement, indicating good internal consistency.(44,45) The inter-rater agreement in pain presence measured by the Kappa statistic was moderate for the "at rest" condition (0.43; 95% CI: 0.16–0.68) and fair for the "with movement" condition (0.25; 95% CI: 0.06–0.47). (46) The ICC was 0.70 (95% CI: 0.33–0.94) for the "at rest" condition and 0.65 (95% CI: 0.38–0.82) for the "with movement" condition, suggesting moderate reproducibility of the CNPI measurements.

Significantly more items were observed with movement compared with the "at rest" condition ($p < .001$ for Rater 1 and for Rater 2, Table 7) supporting the construct validity. Additional evidence supporting construct validity was found in significant associations between the PAS and the CNPI "with movement" conditions ($r^2 = 0.33$, p < .01; $r^2 = 0.41$, $p < .002$). For the CNPI "at rest" condition, neither rater's total CNPI scores was significantly correlated with the PAS. (Table 8)

PAINAD

Table 3 and 5 summarize the PAINAD values among the participants. At rest, the mean scores were 0.2 (SD=0.6) and 0.4 (SD=1.0) for each rater, and with movement means were

1.7 (SD=2.1) and 2.4 (SD=2.1) Similar to the CNPI, PAINAD scores right-skewed with 83–100% being scored a zero at rest and 40–93% with movement.

Cronbach's alphas for total PAINAD at rest were acceptable for rater 2 (.73) and for the total PAINAD with Movement (.70 and .72). However, the Cronbach's alpha for PAINAD "at rest" for Rater 1 was −.04, indicating poor internal consistency.(44,45)

The Kappa statistic was moderate for the PAINAD "with movement" scores (0.54; 95% CI: 0.30–0.72) and fair for the PAINAD "at rest" (0.31; 95% CI: 0.00–0.61) (Table 6).

Construct validity was supported by the significantly higher scores with movement than at rest for both raters (p < .001 for each rater, Table 7). Spearman correlations between the PAINAD and the PAS were statistically significant for the "with movement" (r^2= 0.41, 0.48; *p* < .001) supporting concurrent validity. The correlations between the PAINAD "at rest" condition and the PAS, yielded significant results for one rater (r^2 = 0.27; p = .04) but not for the second rater (r^2 = 0.23; p = .09) (Table 8).

Discussion

The purpose of this study was to examine and compare the reliability and validity of the CNPI and PAINAD in detecting pain in older persons with dementia and known persistent pain. Our findings in this video analysis project suggest that both the CNPI and the PAINAD possess limited validity and reliability, with neither demonstrating clear properties that make it a preferred tool.

Both instruments demonstrated marked floor effects, particularly when participants were at rest. For the CNPI, 80—92% of items were scored as a zero and for the PAINAD, 83—100% of individual items were scored as zero when participants were resting. There were also marked floor effects even with movement. These findings strongly suggest that pain assessment using these two tools should be conducted during or immediately following movement. The original recommendations for scoring of the CNPI include checking "yes" for each of the six categories of behaviors observed with the patient at rest and with movement. The number of checks for both conditions are then totaled, leading to a range of scores from 0—12.. Our findings suggest that the CNPI would be more valuable if it were scored either only on movement, or if the score at rest and movement are compared to determine a change score that reflects pain occurring from movement. Because of the increased user difficulty in creating a change score, we recommend that the CNPI be used similar to the PAINAD with observation during movement only.

Although mean scores on both tools were significantly higher during activity, our study demonstrated marked floor effects even with movement. However, these low scores were consistent with MDS pain data, in which nearly two-thirds of the sample were reported to have "No Pain." The low values also were consistent with the estimates of pain intensity made by nursing assistants who were familiar with the participant; their mean estimate of "usual pain" intensity in the previous week was 3.0 (SD=2.6) on a 0—12 scale. Although participants in the study were selected because they had evidence of moderate pain in the prior week, over 50% reported their pain was intermittent; thus the temporal pain pattern may have influenced observed behaviors as well as NA report. Also, at the time of the videotaping pain could have been controlled or no longer at the level of severity identified for inclusion into the study.

Certain items on both tools contributed markedly to the floor effects during movement. For the PAINAD, the Breathing item was scored > 0 in only 7—10% of cases (Table . This finding is consistent with several other studies.(15,17,19,20) In addition to the floor effects,

the validity of the breathing item as a pain indicator has been questioned.(17,20) This concern also was echoed in the narrative summaries we requested from the raters as they reflected on their experience using both tools. Rater 2 commented that "[the breathing item] was problematic in that especially during activities such as transfer of patients it is not clear whether an increase in the rate of breathing pattern is a pain response or dyspnea due to exertion."

On the CNPI, infrequently seen behaviors included bracing, rubbing, and restlessness (Table 4). Moreover, the change in the frequency of these behaviors between the "at rest" and "with movement" conditions was minimal. Similarly Jones et al (13) reported that restlessness and rubbing were infrequently observed at rest (6% for each behavior) and decreased during movement (4% and 3% respectively). These items may diminish the overall sensitivity of the CNPI to changes in pain, which is consistent with Cohen-Mansfield et al's finding that the CNPI was not able to detect pain treatment effects.(18)

As expected, there were significant differences in "at rest" and "with movement" scores from both raters for both instruments These findings support construct validity by discriminating between a low ("at rest") and a high ("with movement") pain state. The significant associations between the two instruments and agitation for the "with movement" condition provides some evidence for construct validity in that other investigators have documented univariate associations between agitation and pain (34–36,43). However, agitation and behavioral pain measures share some similar items, thereby increasing the likelihood of significant associations. Since agitation may be caused by factors other than physical pain, additional evidence for construct validity is necessary The lack of consistent significant correlations between the tools and agitation for the "at rest" condition strengthens the argument that the PAINAD and CNPI should be used to evaluate pain only during physical activities.

In establishing inter-rater reliability for this study, training was intentionally limited to approximate actual clinical practice. Although training for research purposes generally is extensive so as to achieve very high inter-rater agreement, clinical practice does not allow for this level of practice and comparison of ratings. The 1.5 hour initial training session and follow-up practice session could be accomplished in one to three inservices at facilities. The short training is likely reflected in the inter-rater agreement. While the values of ICC described good agreement between the two raters on the total scores, values of Cohen's kappa pointed at only fair to moderate agreement of the raters on the total presence of pain (total score = 0 vs. >0). This finding casts doubt on the appropriateness of these tools for clinical use. However, the trained RAs rating the videos were not experienced caregivers who were familiar with the residents, and further evaluation with clinical staff is needed for both tools. Additionally, these data suggest that staff must be carefully trained in the use of these tools and that definitions for each behavior be carefully described. Given the typical heavy staff turnover in nursing homes, the findings raise concerns about the ability to maintain a stable cadre of trained personnel, especially CNAs who are adept at using the instruments consistently.

The findings from this tool analysis are limited by the small, homogeneous sample. Because the sample was derived from a subset of available patient data of anl on-going study, no a priori power analyses were conducted. Thus, the apparent lack of differences between the reliability and validity of the instruments may be due to the small sample. Other limitations include the use of only two raters for comparison, and no examination of reliability of rating over time. Future research should examine the tools' performance when administered by staff nurses and/or CNAs as part of their routine care, rather than in the controlled setting of videotaped observations. Additionally, evaluation of these tools' sensitivity in detecting

response to treatment has been limited and is warranted. Further, continued development and testing of other promising pain behavior tools identified in a recent review (See http://prc.coh.org/PAIN-NOA.htm) is warranted.

Conclusion

Detecting pain in persons with dementia is a significant challenge to providers in the LTC environment. Many tools have been developed to assist in this process. Based on the study findings, the CNPI and the PAINAD should be used cautiously both in research and clinical settings. To minimize floor effects, observation of behaviors using these tools should occur during physical activity, and neither instrument should be the sole indicator of pain.; Instead, they should be only one part of a multidimensional pain assessment program that includes more comprehensive screening tools. and other approaches recommended for this population in current practice guidelines and position statements.

Acknowledgments

References

1. Herr K, Coyne PJ, Key T, Manworren R, McCaffery M, Merkel S, et al. Pain assessment in the nonverbal patient: position statement with clinical practice recommendations. Pain Manag Nurs. 2006; 7(2):44–52. [PubMed: 16730317]

2. Hadjistavropoulos T, Herr K, Turk DC, Fine PG, Dworkin RH, Helme R, et al. An interdisciplinary expert consensus statement on assessment of pain in older persons. Clin J Pain. 2007; 23(1 Suppl):S1–43. [PubMed: 17179836]

3. American Geriatrics Society. Pharmacological management of persistent pain in older persons. J Am Geriatr Soc. 2009; 57(8):1331–46. [PubMed: 19573219]

4. Gibson SJ. IASP global year against pain in older persons: highlighting the current status and future perspectives in geriatric pain. Expert Rev Neurother. 2007; 7(6):627–35. [PubMed: 17563246]

5. Herr K, Bjoro K, Decker S. Tools for assessment of pain in nonverbal older adults with dementia: a state-of-the-science review. J Pain Symptom Manage. 2006; 31(2):170–92. [PubMed: 16488350]

6. Herr, KA.; Ersek, M. Measurement of pain and other symptoms in the cognitively impaired. In: Hanks, G.; Cherny, N.; Christakis, N.; Fallon, M.; Kaasa, S.; Portenoy, RK., editors. Oxford Textbook of Palliative Medicine. 4. New York: Oxford University Press; 2009. p. 466-79.

7. Zwakhalen SM, Hamers JP, Abu-Saad HH, Berger MP. Pain in elderly people with severe dementia: a systematic review of behavioural pain assessment tools. BMC Geriatr. 2006; 6:3. [PubMed: 16441889]

8. Stolee P, Hillier LM, Esbaugh J, Bol N, McKellar L, Gauthier N. Instruments for the assessment of pain in older persons with cognitive impairment. J Am Geriatr Soc. 2005; 53(2):319–26. [PubMed: 15673359]

9. van Herk R, van Dijk M, Baar FP, Tibboel D, de Wit R. Observation scales for pain assessment in older adults with cognitive impairments or communication difficulties. Nurs Res. 2007; 56(1):34–43. [PubMed: 17179872]

10. Feldt, KS. Doctoral Dissertation, Dissertation Abstracts International. Vol. 57-09B. University of Minnesota; 1996. Treatment of pain in cognitively impaired versus cognitively intact post hip fractured elders; p. 5574

11. Feldt KS. The checklist of nonverbal pain indicators (CNPI). Pain Manage Nurs. 2000; 1(1):13–21.

12. Feldt KS. Improving assessment and treatment of pain in cognitively impaired nursing home residents. Annals of Long-Term Care. 2000; 8(9):36–41.

13. Jones KR, Fink R, Hutt E, Vojir C, Pepper GA, Scott-Cawiezell J, et al. Measuring pain intensity in nursing home residents. J Pain Symptom Manage. 2005; 30(6):519–27. [PubMed: 16376738]

14. Nygaard HA, Jarland M. The Checklist of Nonverbal Pain Indicators (CNPI): testing of reliability and validity in Norwegian nursing homes. Age Ageing. 2006; 35(1):79–81. [PubMed: 16364939]

15. Warden V, Hurley AC, Volicer L. Development and psychometric evaluation of the Pain Assessment in Advanced Dementia (PAINAD) scale. Journal of the American Medical Directors Association. 2003; 4(1):9–15. [PubMed: 12807591]

16. Costardi D, Rozzini L, Costanzi C, Ghianda D, Franzoni S, Padovani A, et al. The Italian version of the pain assessment in advanced dementia (PAINAD) scale. Arch Gerontol Geriatr. 2007; 44(2):175–80. [PubMed: 16730814]

17. DeWaters T, Faut-Callahan M, McCann JJ, Paice JA, Fogg L, Hollinger-Smith L, et al. Comparison of self-reported pain and the PAINAD scale in hospitalized cognitively impaired and intact older adults after hip fracture surgery. Orthop Nurs. 2008; 27(1):21–8. [PubMed: 18300684]

18. Cohen-Mansfield J, Lipson S. The utility of pain assessment for analgesic use in persons with dementia. Pain. 2008; 134(1–2):16–23. [PubMed: 17481819]

19. Zwakhalen SM, Hamers JP, Berger MP. The psychometric quality and clinical usefulness of three pain assessment tools for elderly people with dementia. Pain. 2006; 126(1–3):210–20. [PubMed: 16890355]

20. van Iersel T, Timmerman D, Mullie A. Introduction of a pain scale for palliative care patients with cognitive impairment. Int J Palliat Nurs. 2006; 12(2):54–9. [PubMed: 16603993]

21. Schuler MS, Becker S, Kaspar R, Nikolaus T, Kruse A, Basler HD. Psychometric properties of the German "Pain Assessment in Advanced Dementia Scale" (PAINAD-G) in nursing home residents. J Am Med Dir Assoc. 2007; 8(6):388–95. [PubMed: 17619037]

22. Leong IY, Chong MS, Gibson SJ. The use of a self-reported pain measure, a nurse-reported pain measure and the PAINAD in nursing home residents with moderate and severe dementia: a validation study. Age Ageing. 2006; 35(3):252–6. [PubMed: 16497681]

23. Lane P, Kuntupis M, MacDonald S, McCarthy P, Panke JA, Warden V, et al. A pain assessment tool for people with advanced Alzheimer's and other progressive dementias. Home Healthc Nurse. 2003; 21(1):32–7. [PubMed: 12544460]

24. Hutchison RW, Tucker WF Jr, Kim S, Gilder R. Evaluation of a behavioral assessment tool for the individual unable to self-report pain. Am J Hosp Palliat Care. 2006; 23(4):328–31. [PubMed: 17060298]

25. Herr, K.; Bursch, H.; Black, B. State of the Art Review of Tools for Assessment of Pain in Nonverbal Older Adults. 2008. Available from: http://prc.coh.org/PAIN-NOA.htm

26. Herr K, Bursch H, Ersek M, Miller L, Swafford K. Behavioral assessment tools for pain assessment in nursing homes: Expert consensus recommendations for practice. Journal of Gerontological Nursing. In press.

27. Centers For Medicare & Medicaid Services. Revised Long-Term Care Facility Resident Assessment Instrument User's Manual, Version 2.02008. July 29. 2009 Available from: http://www.cms.hhs.gov/nursinghomequalityinits/20_NHQIMDS20.asp

28. Morris JN, Fries BE, Mehr DR, Hawes C, Phillips C, Mor V, et al. MDS Cognitive Performance Scale. Journal of Gerontology. 1994; 49(4):M174–82. [PubMed: 8014392]

29. Herr K, Spratt KF, Garand L, Li L. Evaluation of the Iowa pain thermometer and other selected pain intensity scales in younger and older adult cohorts using controlled clinical pain: a preliminary study. Pain Med. 2007; 8(7):585–600. [PubMed: 17883743]

30. Taylor L, Harris J, Epps C, Herr K. Psychometric Evaluation of Selected Pain Intensity Scales for Use with Cognitively Impaired and Cognitively Intact Older Adults. Rehabilitation Nursing. 2005; 30(2):55–61. [PubMed: 15789697]

31. Ware LJ, Epps CD, Herr K, Packard A. Evaluation of the Revised Faces Pain Scale, Verbal Descriptor Scale, Numeric Rating Scale, and Iowa Pain Thermometer in older minority adults. Pain Manag Nurs. 2006; 7(3):117–25. [PubMed: 16931417]

32. Taylor LJ, Herr K. Pain intensity assessment: a comparison of selected pain intensity scales for use in cognitively intact and cognitively impaired African American older adults. Pain Manag Nurs. 2003; 4(2):87–95. [PubMed: 12836153]

33. Lynn Snow A, Cook KF, Lin PS, Morgan RO, Magaziner J. Proxies and other external raters: methodological considerations. Health Serv Res. 2005; 40(5 Pt 2):1676–93. [PubMed: 16179002]

34. Buffum MD, Miaskowski C, Sands L, Brod M. A pilot study of the relationship between discomfort and agitation in patients with dementia. Geriatric Nursing. 2001; 22(2):80–5. [PubMed: 11326214]

35. Zieber CG, Hagen B, Armstrong-Esther C, Aho M. Pain and agitation in long-term care residents with dementia: use of the Pittsburgh Agitation Scale. Int J Palliat Nurs. 2005; 11(2):71–8. [PubMed: 15798498]

36. Manfredi PL, Breuer B, Wallenstein S, Stegmann M, Bottomley G, Libow L. Opioid treatment for agitation in patients with advanced dementia. Int J Geriatr Psychiatry. 2003; 18(8):700–5. [PubMed: 12891637]

37. Rosen J, Burgio L, Kollar M, Cain M, Allison M, Fogleman M, et al. The Pittsburgh Agitation Scale: A user-friendly instrument for rating agitation in dementia patients. Am J Geriatr Psychiatry. 1994; 2(1):52–9. [PubMed: 21629007]

38. Richards JS, Nepomuceno C, Riles M, Suer Z. Assessing pain behavior: the UAB Pain Behavior Scale. Pain. 1982; 14(4):393–8. [PubMed: 7162841]

39. Hurley AC, Volicer BJ, Hanrahan PA, Houde S, Volicer L. Assessment of discomfort in advanced Alzheimer patients. Research in Nursing & Health. 1992; 15(5):369–77. [PubMed: 1529121]

40. Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S. The FLACC: a behavioral scale for scoring postoperative pain in young children. Pediatr Nurs. 1997; 23(3):293–7. [PubMed: 9220806]

41. Davison, A.; Hinkley, D. Bootstrap methods and their application. Cambridge, United Kingdom: Cambridge University Press; 1997.

42. Garson, G. Validity. Statnotes: Topics in Multivariate Analysis. 2008. Available from: http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm.

43. Villanueva MR, Smith TL, Erickson JS, Lee AC, Singer CM. Pain Assessment for the Dementing Elderly (PADE): reliability and validity of a new measure. J Am Med Dir Assoc. 2003; 4(1):1–8. [PubMed: 12807590]

44. Polit, DF.; Beck, CT. Nursing research: principles and methods. 7. Philadelphia: Lippincott Williams & Wilkins; 2004.

45. Tabachnick, B.; Fidell, L. Using multivariate statistics. 4. Boston: Allyn and Bacon; 2000.

46. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33(1):159–74. [PubMed: 843571]

Table 1

Sample Demographic and Pain Characteristics (n = 60)

	Percent	Mean (SD)
Age (Range: 69—103 years)		89.0 (6.8)
Female	88.1	
Marital status		
Widowed	61.0	
Married	27.1	
Divorced	6.8	
Never married	5.1	
Ethnicity		
White, not Hispanic	93.2	
Black, not Hispanic	3.4	
Asian/Pacific Islander	3.4	
Education: High School or less	67.8	
Cognitive Performance Scale [*](0—6)		3.9 (1.2)
Pain Frequency from the MDS		
No pain	61.0	
Pain less than daily	23.7	
Pain daily	15.3	
Pain Intensity from the MDS (n = 23)		
Mild pain	52.2	
Moderate pain	34.8	
Excruciating pain	13.0	
Number of Documented Painful Diagnoses		1.5(1.1)
0	18.6	
1	42.4	
2	20.3	
3	11.9	
4	6.8	
Diagnosis of Osteoarthritis	62.7	
CAN report of "usual pain" during the previous week[#] (0—12 scale)		3.1(2.6)
Pittsburgh Agitation Scale[‡] (0—16 scale)		3.5(3.4)

MDS = Minimum Data Set; CNA = Certified Nursing Assistant

[*] Higher scores indicating greater cognitive impairment

[#] Higher scores indicate greater pain intensity

[‡] Higher scores indicate greater agitation

Table 2

CNPI scores* by Rater and Activity Level

Rater		N	CNPI Mean (SD)	Median	Observations when CNPI = 0 N (%)	Cronbach's Alpha
RA 1	Rest	60	0.9 (2.0)	0	46 (76.7)	.97
	Movement	60	1.9 (2.2)	1	25 (41.7)	.90
RA 2	Rest	59	0.8 (1.7)	0	42 (71.2)	.92
	Movement	60	2.0 (1.6)	2	9 (15.0)	.74

RA = Research Assistant; CNPI = Checklist of Nonverbal Pain Indicators

*
Range of scores: 0–6 for each condition, i.e., Rest and Movement, higher scores may indicate greater pain

Table 3

PAINAD Scores[*] by Rater and Activity Level

Rater		N	Mean (SD)	Median	Observations when PAINAD=0 N (%)	Cronbach's alpha
RA 1	Rest	59	0.2 (0.6)	0	50 (84.7)	−.04
	Movement	59	1.7 (2.1)	0	26 (44.1)	.72
RA 2	Rest	59	0.4 (1.0)	1	47 (79.7)	.73
	Movement	60	2.4 (2.1)	3	19 (31.7)	.70

RA = Research Assistant; PAINAD = Pain Assessment in Advanced Dementia

[*]
Range of scores: 0–12, higher scores indicate greater pain

Table 4

Total CNPI Indicators* for both Raters at Rest and with Movement

	Present At Rest		Present With Movement	
	Rater 1	**Rater 2**	**Rater 1**	**Rater 2**
Vocalizations	18%	17%	40%	52%
Grimaces	20%	17%	50%	73%
Bracing	13%	8%	20%	12%
Rubbing	13%	10%	20%	12%
Restlessness	15%	18%	18%	15%
Verbal complaint	13%	10%	40%	37%

CNPI = Checklist of Nonverbal Pain Indicators

Table 5

PAINAD Indicators at Rest and with Movement

	Scoring 1 or higher At Rest		Scoring 1 or higher With Movement	
	Rater 1	**Rater2**	**Rater 1**	**Rater 2**
Breathing	7%	2%	10%	7%
Negative vocalization	7%	7%	41%	47%
Facial expression	3%	17%	41%	60%
Body language	2%	10%	14%	25%
Consolability	0%	5%	24%	42%

PAINAD = Pain Assessment in Advanced Dementia

Table 6

Reproducibility for PAINAD and CNPI

	N	ICC for total score (95% CI[*])	Kappa for pain presence: total score =0 vs. >0 (95% CI[*])
CNPI REST	59	0.70 (0.33–0.94)	0.43 (0.16–0.68)
CNPI MOVEMENT	60	0.65 (0.38–0.82)	0.25 (0.06–0.47)
PAINAD REST	58	0.24 (0.02–0.58)	0.31 (0.00–0.61)
PAINAD MOVEMENT	59	0.80 (0.65–0.89)	0.54 (0.30–0.72)

[*] adjusted bootstrap percentile confidence interval; PAINAD = Pain Assessment in Advanced Dementia; CNPI = Checklist of Nonverbal Pain Indicators

Table 7

Difference in mean CNPI and PAINAD scores at rest and during movement

		Mean difference (SD)	P value[*]
CNPI	Rater 1	0.95 (1.7)	<0.001
	Rater 2	1.2 (1.2)	<0.001
PAINAD	Rater 1	1.5 (2.1)	<0.001
	Rater 2	2.0 (1.9)	<0.001

[*] paired t-test; PAINAD = Pain Assessment in Advanced Dementia; CNPI = Checklist of Nonverbal Pain Indicators

TABLE 8

Spearman Correlations between Pittsburgh Agitation Scale and the CNPI and PAINAD Scores

	Pittsburgh Agitation scores
CNPI at Rest (Rater 1)	.17
CNPI at Rest (Rater 2)	.16
CNPI with Movement (Rater 1)	.33[*]
CNPI with Movement (Rater 2)	.41[*]
PAINAD at Rest (Rater 1)	.27[*]
PAINAD at Rest (Rater 2)	.23
PAINAD with Movement (Rater 1)	.41[**]
PAINAD with Movement (Rater 2)	.48[**]

[*] $P < .05$;

[**] $P < .001$;

PAINAD = Pain Assessment in Advanced Dementia; CNPI = Checklist of Nonverbal Pain Indicators