Pain is an important problem in the intensive care unit (ICU), and inadequate pain assessment and management have been linked to increased morbidity and mortality. The physiological response to pain is almost universally adverse, causing potentially fatal unstable hemodynamic status, alterations in immune system functioning, hyperglycemia, and increased release of catecholamine, cortisol, and antidiuretic hormones. Moreover, uncontrolled pain has been implicated in a variety of psychosocial effects, including depression, anxiety, delirium, posttraumatic stress disorder, and disorientation. Despite the acknowledgment that pain is a common stressor in the ICU, high rates of uncontrolled pain in critically ill patients remain common. This situation can be attributed, in part, to circumstances, such as mechanical ventilation or unstable hemodynamic status, that preclude the assessment of pain by self-report. Despite strong evidence that documentation of pain assessment improves pain management and decreases patients’ pain, no pain assessment instrument has been universally recommended for use in critically ill patients incapable of self-reporting.
Background and Significance

The pathophysiology of pain is complex, but involves the release of inflammatory mediators after tissue injury. The mediators stimulate nociceptors, which transmit pain impulses to the dorsal horn of the spinal cord. The pain impulses arrive at the somatosensory cortex, where the localization and meaning of pain occur. Pain can be classified as acute or chronic, depending on duration, and as peripheral or central, depending on location. Further classifications are used to describe the pain source, such as from injury to the skin (cutaneous), nerves (neuropathic), muscles and bones (somatic), or organs (visceral).

Nearly 5 million patients are admitted to the ICU each year, and an estimated 71% of those patients remember experiencing pain during their stay. Painful procedures, such as turning and tracheal suctioning, are common in the ICU and precipitate acute pain. In addition, many critically ill patients have a history of chronic pain, which complicates assessment and treatment. Pain is one of the most common symptoms in critically ill patients and is experienced by each patient in a unique manner.

The routine use of an appropriate assessment of pain has been mandated by the Agency for Healthcare Research and Quality and the Joint Commission. Professional organizations such as the American Association of Critical-Care Nurses, the American College of Chest Physicians, the Society for Critical Care Medicine, and the American Society for Pain Management agree. All of these organizations advocate for implementation of standardized pain assessment tools that include behavioral indicators in patients who are sedated and receiving mechanical ventilation and incapable of self-reporting or whose self-reports may be unreliable. Furthermore, studies have indicated that appropriate treatment of pain and anxiety is associated with decreased duration of mechanical ventilation (8 vs 11 days; \( P < .01 \)), decreased rate of nosocomial infections, and increased patient satisfaction with pain control. Additionally, systematic pain assessment in critically ill patients can decrease ICU length of stay (13 vs 18 days; \( P < .01 \)).

OVID, PubMed, Medline, and CINAHL databases were used to search the literature. The key words pain, assessment, tool, instrument, scale, intensive care, critical care, and critically ill were used. Articles published between 2000 and 2010 and available in English were included in the review. Additional inclusion criteria included publication in a professional, scholarly, peer-reviewed nursing and health care journal. Articles that focused on infants and children, addressed pain assessment outside the ICU, or focused on the evaluation of chronic pain were excluded. A total of 1120 articles were found; of these, 17 met the inclusion criteria. A few studies are included that did not meet every inclusion criterion but did contain important information.

Review of the Literature

The review focuses on tools developed for and tested in adult critical care patients. These instruments include the Pain Assessment and Intervention Notation (PAIN) algorithm, the Nonverbal Pain Assessment Tool (NPAT), the Adult Nonverbal Pain Scale (NVPS), the Behavioral Pain Scale (BPS), and the Critical-Care Pain Observation Tool (CPOT).

PAIN Algorithm

The PAIN algorithm, developed in 2001, consists of 3 parts: pain assessment, the ability of the patient to tolerate opioids, and guidelines for analgesic treatment decisions and documentation. The pain assessment section of the tool contains both behavioral (movement, facial cues, posturing) and physiological (increased heart rate, respiratory rate, and blood pressure and perspiration or pallor) dimensions.

The initial testing of the tool included a study sample of 11 nurses who used the instrument to assess 31 postoperative patients in the ICU or postanesthesia care unit. The algorithm was helpful in providing a “systematic approach” to pain assessment and guidance of analgesic administration. However, 4 of the 11 nurses thought that the PAIN tool was not helpful: it was “too long and cumbersome” to be used in “everyday practice.” Reliability
and validity of the tool were not determined. Although the researchers found that the instrument could be a useful training technique for beginning ICU nurses, the length of the tool has limited its clinical usefulness, and no further testing has been done.\textsuperscript{17}

**Nonverbal Pain Assessment Tool**

The NPAT consists of 5 domains: emotion, movement, verbal cues, facial cues, and positioning/guarding.\textsuperscript{10} Two separate scoring systems are provided on the instrument for use in both verbal and nonverbal patients. Scores range from 0 to 10 points, with higher numbers indicating higher severity of pain. The initial testing of the NPAT consisted of 3 phases of validity testing and 2 revisions of the tool. In the third phase of validity testing, the investigators analyzed the correlation between the NPAT score and the patient's self-reported pain score in a sample of 50 general medical-surgical, postoperative patients.

Interrater reliability was high (concordance coefficient, 0.72; 95% confidence interval), and criterion validity (Table 1) was moderately strong when the NPAT was compared with the gold standard of self-report (concordance coefficient, 0.66; 95% confidence interval).\textsuperscript{19} Although designed for noncommunicative critically ill patients, the NPAT has not been validated in this population of patients; the initial testing of the third phase was done in verbal medical-surgical patients. In addition, the inclusion of a “verbal cues” domain in a nonverbal pain assessment tool is perplexing. Previous versions of the NPAT studied in the ICU showed weak validity (concordance correlation coefficient, 0.31; 95% confidence interval and concordance correlation coefficient, 0.21; 95% confidence interval, respectively). Since the initial testing in 2010, no further reliability or validity testing of the NPAT has been completed.\textsuperscript{19}

**Adult Nonverbal Pain Scale**

The NVPS (Table 2) was based on the Faces, Legs, Activity, Cry, Consolability (FLACC) scale.\textsuperscript{20} Like the PAIN algorithm and the NPAT, the NVPS contains behavioral dimensions (facial expression, activity, and guarding) and physiological dimensions (heart rate, blood pressure, and respiratory rate) that are graded in severity. The final domain of the NVPS includes autonomic indicators such as dilated pupils, diaphoresis, flushing, or pallor. Each domain is ranked from 0 to 2, with a total score between 0 (no pain) and 10 (maximum pain).

The initial testing of the NVPS was conducted in a burn trauma unit on 59 patients, with a total of 100 paired assessments performed. Internal consistency was high (\(\alpha = .78; P < .001\)). The authors\textsuperscript{20} concluded that the NVPS was a valid observational pain scale in this patient population because correlation of the NVPS scores with the FLACC scores was high (\(\alpha = .86; P = .05\)). However, use of the FLACC as a gold standard in this study cannot be supported; the FLACC was developed for use in children and has not been validated in adults.
In a study by Kabes et al., a revised version of the NVPS was compared with the original NVPS for reliability and validity testing. The revised NVPS includes a new "respiratory" category that replaces the physiological II dimension of the original scale. The dimension includes an assessment of the amount of deviation from the baseline respiratory rate, as well as oxygen saturation as measured by pulse oximetry and level of compliance with the ventilator. Using a nonexperimental design, nurses in a trauma-surgical ICU assessed patients before, during, and at rest after a painful nursing procedure.

The levels of intrarater reliability (≥90% in 94.7% with original NVPS and 90.8% with the revised NVPS) and internal consistency (Cronbach α = 0.36 before, 0.72 during, 0.71 after) were high. Both the original and the revised NVPS showed significant differences between ratings before, during, and after the painful procedure (original, 135.86; P < .001; n = 121; revised, 145.05; P < .001; n = 121). Correlations between NVPS scores and self-report scores were not determined.

Marmo and Fowler used a repeated-measures study design to examine the validity of the NVPS in a sample of 25 critically ill patients after open heart surgery. The NVPS, CPOT, and FLACC were administered for a total of 300 paired, independent observations before, during, and after a painful procedure (endotracheal suctioning, turning). The NVPS was reliable, with a Cronbach α coefficient of 0.89, but disagreement was high between nurse raters in analyses of the facial expression dimension (25% of the total observations).

Wibbenmeyer et al did a similar study in a sample of 38 burn patients. A total of 225 paired assessments were completed by nursing staff who were "briefly educated" on use of the CPOT and the NVPS. No description of the education was provided. The assessments were completed with the patients at rest and again shortly after a painful stimulus (wound care, physical or occupational therapy). Because all of the patients were able to respond verbally, a self-report of pain intensity was also obtained. The NVPS had good internal consistency (Cronbach α = 0.80), but intrarater reliability was merely fair (Pearson correlation coefficient, 0.59). The poor reliability may have been due to the limited education of the data collectors on the appropriate administration of the tool. Discriminate validity was indicated by the change in mean NVPS scores beginning with a mean at-rest score of 0.19 and increasing to 0.44 (P < .001) after the noxious stimulus was applied. Correlation between the NVPS and the Numerical Rating Scale (NRS) was poor (Pearson correlation coefficient, 0.38; P < .01), indicating that the NVPS is a poor indicator of pain intensity.

Finally, Topolovec-Vranic et al. found that implementation of the NVPS was associated with a trend toward decreased levels of severe pain (55% vs 35% of patients reporting severe pain), although these findings were not statistically significant, and increased frequency of pain assessment (29% vs 59%; P < .001). The investigators acknowledged that the trend may have been due to an increased awareness of the need to assess pain after the

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Adult Nonverbal Pain Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Categories</strong></td>
<td><strong>0</strong></td>
</tr>
<tr>
<td><strong>Face</strong></td>
<td>No particular expression or smile</td>
</tr>
<tr>
<td><strong>Activity (movement)</strong></td>
<td>Lying quietly, normal position</td>
</tr>
<tr>
<td><strong>Guarding</strong></td>
<td>Lying quietly, no positioning of hands over areas of body</td>
</tr>
<tr>
<td><strong>Physiology (vital signs)</strong></td>
<td>Stable vital signs</td>
</tr>
<tr>
<td><strong>Respiratory</strong></td>
<td>Baseline RR/SpO₂ Compliant with ventilator</td>
</tr>
</tbody>
</table>

Abbreviations: HR, heart rate; RR, respiratory rate; SBP, systolic blood pressure; SpO₂, pulse oximetry. Instructions: Each of the 5 categories is scored from 0-2, which results in a total score between 0 and 10.
NVPS education sessions that were provided to the staff before the data collection period.

**Behavioral Pain Scale**

The BPS\(^{26}\) (Table 3) was based on the work of Puntillo et al\(^{27}\) that identified unique behaviors present in patients undergoing a noxious stimulus. The BPS is composed of 3 observational items (facial expression, upper limbs, and compliance with ventilation) that are scored from 1 to 4, with higher numbers indicating higher levels of discomfort. The total BPS score can range from 3 (no pain) to 12 (most pain). The BPS was initially tested by using a quasi-experimental design and a sample of 30 patients receiving mechanical ventilation.\(^{26}\) The BPS scores were assessed at rest and then during a nonpainful procedure (application of compression stockings or dressing change of a central venous catheter) or during a nociceptive procedure (endotracheal suctioning or turning). Patients who had painful procedures had higher BPS scores (range, 4.6-5.2) than did patients who had nonpainful procedures (BPS score range, 3.3-3.7). A test-retest procedure was then performed in 31 cases, resulting in BPS scores that increased from 3.0 to 3.3 at rest to 4.0 to 4.8 during a painful procedure, establishing construct validity. Interrater agreement was high (\(\alpha = 0.94\)), and 24 of the 28 evaluators reported that the BPS was easy to use and took a mean of 2 to 5 minutes to complete.

Aïssaoui et al\(^{28}\) tested the BPS in 30 general ICU patients in a total of 360 observations. All patients included in the study were sedated and were receiving mechanical ventilation. In paired, independent observations, BPS scores were obtained at rest and during painful procedures (endotracheal suctioning or peripheral vein cannulation). The BPS had good internal consistency (Cronbach \(\alpha = 0.72\)), and interrater agreement was high (intraclass correlation coefficient, 0.95). Validity was established by the change in BPS scores (mean, 3.9; SD, 1.1 at rest and mean, 6.8; SD, 1.9 during procedures). The elevated BPS scores at rest are interesting and most likely support hypotheses provided in previous research that critically ill patients may have moderate levels of pain even at rest.\(^{29}\)

In another study, Young et al\(^{30}\) used a descriptive, repeated-measures design to analyze the reliability and validity of the BPS. A total of 44 patients receiving mechanical ventilation whose hemodynamic status was stable were assessed before and after a nonpainful (eye care) and a painful (repositioning) procedure. The mean BPS score increased significantly after the repositioning procedure (score, 3.36 before vs 5.02 after). In contrast, the mean BPS score did not change significantly after the eye care procedure (score, 3.23 before vs 3.38 after), supporting discriminate validity of the BPS. However, interrater agreement varied widely, with good agreement (82%-91%) during rest states, but a marked decrease in agreement after a painful procedure (36%-46%). These results indicate that raters more easily agreed on the level of pain when the level was low and that variance between raters occurred when pain levels were high.

Ahlers et al\(^{31}\) obtained 371 independent determinations of the BPS and NRS (when possible) in 113 critically ill patients. The BPS was completed for all patients receiving mechanical ventilation; the NRS was either estimated by the bedside nurse or obtained from the patient if the patient was able to respond. Having a bedside nurse estimate the NRS score is a questionable practice, because this study\(^{31}\) indicated that the NRS score as reported by the patient correlated with the score estimated by the

---

**Table 3  Behavioral Pain Scale**

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Score(^{a})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facial expression</td>
<td>Relaxed</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Partially tightened (eg, brow lowering)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Fully tightened (eg, eyelid closing)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Grimacing</td>
<td>4</td>
</tr>
<tr>
<td>Upper limb movements</td>
<td>No movement</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Partially bent</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Fully bent with finger flexion</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Permanently retracted</td>
<td>4</td>
</tr>
<tr>
<td>Compliance with mechanical</td>
<td>Tolerating movement</td>
<td>1</td>
</tr>
<tr>
<td>ventilation</td>
<td>Coughing but tolerating ventilation for most of the time</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Fighting ventilator</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Unable to control ventilation</td>
<td>4</td>
</tr>
</tbody>
</table>

\(^{a}\) Score ranges from 3 (no pain) to 12 (maximum pain).
nurse only 73% of the time. Despite the problematic study design, the interrater reliability of the BPS was good ($\kappa = 0.67$; 95% CI, 0.54-0.80). Of note, NRS scores were 0 in only 5% of the observations, whereas BPS scores were 3, indicating no pain, in 68%. These data illuminate an interesting question in the study by Payen et al\textsuperscript{26} in which the proportion of low BPS scores at rest was high. Possibly a BPS score of 3 is not truly representative of a pain-free state as previously postulated. In addition, the BPS equates lack of body movement with a pain-free state, which conflicts with research\textsuperscript{32} that indicates a correlation between decreased movement and increased pain levels.

In a follow-up study, Ahlers et al\textsuperscript{33} examined the reliability and validity of the BPS in both deeply sedated patients and patients undergoing conscious sedation for a painful procedure. Nurses made 175 observations in a convenience sample of 80 medical-surgical ICU patients at 4 times: at rest, during a nonpainful procedure, during a painful procedure, and after the painful procedure. The assessments were done independently, and communicative patients were asked to provide a report of their pain intensity by using the Verbal Rating Scale\textsuperscript{429} if able to do so after the BPS scoring was completed. Interrater reliability of the BPS was excellent ($\kappa = 0.80$ for deeply sedated patients and 0.83 for conscious sedated patients). Internal consistency was moderate ($\alpha = 0.63$ in deeply sedated patients and 0.66 in conscious sedated patients). BPS scores were higher during painful procedures than at rest in both deeply sedated patients (mean, 5.1; 95% CI, 4.8-5.5 vs mean, 3.4; 95% CI, 3.3-3.5) and conscious sedated patients (mean, 5.4; 95% CI, 4.9-5.9 vs mean, 3.8; 95% CI, 3.5-4.1), thus establishing construct validity. The BPS and the Verbal Rating Scale\textsuperscript{4} had a strong positive correlation during painful procedures ($r = 0.67$; $P < .001$), indicating that the BPS is a valid tool to be used in both deeply sedated and moderately sedated patients.

### Critical-Care Pain Observation Tool

The CPOT\textsuperscript{34} (Table 4) was developed on the basis of retrospective chart reviews to determine common pain notations and findings with focus groups of ICU clinicians. This instrument is designed for use in both intubated and nonintubated critical care patients. Four domains—facial expressions, movements, muscle tension, and ventilator compliance—are scored from 0 to 2; total scores range from 0 (no pain) to 8 (most pain). The CPOT contains

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facial expression</td>
<td>No muscular tension observed</td>
<td>Relaxed, neutral 0</td>
</tr>
<tr>
<td>Presence of frowning, brow lowering, orbit tightening, and levator contraction All of the above facial movements plus eyelid tightly closed</td>
<td>Tense 1</td>
<td></td>
</tr>
<tr>
<td>Body movements</td>
<td>Does not move at all (does not necessarily mean absence of pain) Slow, cautious movements, touching or rubbing the pain site, seeking attention through movements Pulling tube, attempting to sit up, moving limbs/thrashing, not following commands, striking at staff, trying to climb out of bed</td>
<td>Absence of movements 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Protection 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Restlessness 2</td>
</tr>
<tr>
<td>Muscle tension</td>
<td>No resistance to passive movements</td>
<td>Relaxed 0</td>
</tr>
<tr>
<td>Evaluation by passive flexion and extension of upper extremities</td>
<td>Resistance to passive movements</td>
<td>Tense, rigid 1</td>
</tr>
<tr>
<td></td>
<td>Strong resistance to passive movements, inability to complete them</td>
<td>Very tense or rigid 2</td>
</tr>
<tr>
<td>Compliance with the ventilator (intubated patients)</td>
<td>Alarms not activated, easy ventilation</td>
<td>Tolerating ventilator or movement 0</td>
</tr>
<tr>
<td>Alarms stop spontaneously</td>
<td>Fighting ventilator 2</td>
<td></td>
</tr>
<tr>
<td>Asynchrony: blocking ventilation, alarms frequently activated</td>
<td>Coughing but tolerating 1</td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocalization (extubated patients)</td>
<td>Talking in normal tone or no sound</td>
<td>Talking in normal tone or no sound 0</td>
</tr>
<tr>
<td>Sighing, moaning</td>
<td>Sighing, moaning 1</td>
<td></td>
</tr>
<tr>
<td>Crying out, sobbing</td>
<td>Crying out, sobbing 2</td>
<td></td>
</tr>
<tr>
<td>Total, range</td>
<td></td>
<td>0-8</td>
</tr>
</tbody>
</table>
evidence-based descriptors that are operationally defined for each domain, and the content validity index of all indicators was 0.88 to 1.0, according to an analysis of the results of a questionnaire provided to physicians and critical care nurses. The CPOT was originally developed in French and was tested in a convenience sample of 105 cardiac surgery patients.

Gélinas et al\textsuperscript{36} used a repeated-measures design and trained data collectors to obtain the CPOT score of 105 patients at 3 times: at rest, immediately after repositioning (nociceptive procedure), and at recovery (20 minutes after repositioning). After completion of the assessment by 2 observers, the patient was asked to indicate the presence or absence of pain by nodding the head yes or no. Patients who had delirium (as indicated by the Confusion Assessment Method for the Intensive Care Unit) were excluded from the study. Interrater reliability was moderate to high (weighted \(\kappa\) coefficient = 0.52-0.88) when tested between the same 2 data collectors. Criterior validity was established by comparing the patient’s self-report of pain and the concurrent CPOT score. In patients reporting pain, the mean CPOT score was 1.62 to 3.65; in patients reporting no pain, the mean score was 0.49 to 2.11. Although the mean CPOT scores were lower in the patients reporting no pain, the CPOT appeared to provide an overestimation of pain in some instances, especially in the immediate period after a nociceptive procedure (mean, 2.11; SD, 0.90). Discriminate validity was supported by the finding that the CPOT scores were significantly higher \(P < .001\) during positioning \((t = -15.96)\) than at rest \((t = -9.01)\). Although the study sample was large, interrater reliability and the homogeneity of the sample were acknowledged limitations.

In a post hoc data analysis, Gélinas et al\textsuperscript{36} evaluated the sensitivity and specificity of the CPOT. They found that the tool had a sensitivity of 86% and a specificity of 78% during painful procedures. However, the sensitivity decreased to 83% before a painful procedure and was 63% after the procedure. Specificity remained high at 83% and 97%, respectively. A cross-over observational design was used to test the first English version of the CPOT in a group of 30 conscious and 25 unconscious general ICU patients.\textsuperscript{37} Additional data were collected in this study,\textsuperscript{37} including physiological data and patients’ self-report of pain via use of the Faces Pain Thermometer. Interrater reliability was acceptable (intraclass correlation coefficients, 0.80-0.93) among the 51 nurses who collected the data. Discriminate validity was established by detection of elevations in heart rate and blood pressure that occurred in accordance with elevated CPOT scores. Yet, physical motion is known to increase heart rate and blood pressure to compensate for increased oxygen demand, making this method of establishing validity problematic. However, when patients’ self-reported pain values were compared with the observer-derived CPOT scores, the positive predictive value of the CPOT was high (85.7%). All of the nurse respondents reported that the CPOT directives were clear and that the tool was easy to use. The majority of the nurses also reported that the CPOT was quick to use (78%) and that they would recommend use of the CPOT routinely in practice (72.7%).

Marmo and Fowler\textsuperscript{23} tested the CPOT in patients after heart surgery and found that the tool had high reliability \((\alpha = 0.89)\). These researchers\textsuperscript{23} were also the first to report the internal consistency of the CPOT (56%-100% agreement). The CPOT was also included in the previously discussed study by Wibbenmeyer et al,\textsuperscript{24} who reported a high internal consistency \((\text{Cronbach } \alpha, 0.71)\) and good discriminate validity \((\text{mean scale scores } = 0.27 \text{ at rest to } 0.56 \text{ after noxious stimulation})\). As with the NVPS, Wibbenmeyer et al reported that the interrater reliability of the CPOT was poor \((\text{Pearson correlation coefficient, 0.63; } P < .001)\). However, this poor reliability could again be due to the limited amount of training the assessors received before data collection. In the study by Marmo and Fowler,\textsuperscript{23} an 85% agreement between raters was ensured before the study began. In the study by Wibbenmeyer et al,\textsuperscript{24} interrater agreement was not optimized before data collection.

Vázquez et al\textsuperscript{39} conducted a prospective, repeated-measures study in a 12-bed general ICU in Spain. A total of 330 paired observations were completed in a study sample of 96 critically ill patients. Observations were conducted before, during, and after a repositioning procedure. Interrater reliability of the CPOT was excellent \((\kappa = 0.79)\), and discriminate validity was good; mean scores were 0.27 (SD, 0.64) at rest and 1.93 (SD, 1.41) during the procedure.

Finally, Gélinas et al\textsuperscript{40} examined the effects of implementing the CPOT in a general ICU in Canada. Before implementation, interrater agreement for patients at rest was associated with decreased frequency of administration of sedatives and analgesics.
was high (95%-100%), and agreement during turning was acceptable (73%-91%). After implementation of the CPOT, interrater agreement improved to 86% to 100%. When pain assessment practices were analyzed by using descriptive statistics, reports of pain assessments were 3 to 4 times more frequent after implementation than they were before implementation. Interestingly, implementation of the CPOT was associated with decreased frequency of administration of sedatives and analgesics. Gelinas et al40 provide 2 possible explanations: increased ability of nursing staff to discern pain from other symptoms (such as anxiety) or decreased number of trauma patients in the group after implementation because of a change in the center’s trauma designation.

Analysis of the Literature

The published evidence indicates that more research to find the ideal instrument to assess pain in critically ill nonverbal adults is needed. Although multiple tools are available, few have been reliable and valid across a multitude of patient populations and settings (Table 5). However, some tools clearly have been tested much more than others have. The PAIN algorithm and NPAT have each been used in only a single study, and the findings in both studies were of limited value. The original testing of the NVPS is of limited value because of the study’s nonexperimental design and the use of the FLACC scale as the gold standard for comparison. Although the NVPS is the only tool that includes dimensions of physiological data, these indicators have been some of the least sensitive markers for the presence of pain.41

An overwhelming majority of the studies provide support for the reliability and validity of both the BPS and the CPOT in detecting pain in noncommunicative critically ill adults. The internal consistency and interrater reliability of both the BPS and the CPOT have been shown in more than 1 prospective study. However, comparisons of study findings indicate that the CPOT has been more internally consistent than the BPS (Cronbach $\alpha=0.71$-$0.89$ vs Cronbach $\alpha=0.63$-$0.66$, respectively). Whether these data translate into a true implication for practice remains to be seen; high levels of internal consistency may simply indicate redundancy within the tool.41 Moreover, the range of statistical analyses used in the data analysis for the CPOT was broader than that used for the BPS.

The results of construct, criterion, and discriminate validity analyses in a variety of patient populations have been favorable for both the BPS and the CPOT; both tools indicated a statistically significant increase in scores after a painful procedure was performed. However, the studies by Payen et al26 and Young et al30 also indicated an increase in the BPS score after a nonnociceptive procedure, making the specificity of the instrument a concern. Although the sensitivity and specificity of the CPOT have been investigated,32 no data on sensitivity or specificity of the BPS have been reported. An interesting finding is the mean at-rest scores, which were elevated with both tools (BPS score, 3.7; CPOT score, 0.27). These findings suggest baseline pain is present in a majority of critically ill patients, consistent with the results of a previous study29 on characteristics of pain in the ICU.

An important finding of the validity testing of both tools is a lack of strong correlation of the observational pain tools with patients’ self-reports. The correlation was especially poor when patients were scored as having no pain (BPS score, 3; CPOT score, 0), and only a few patients had no pain on self-report scales. These findings suggest that although scores increased during nociceptive procedures, indicating the presence of increased pain, the BPS and the CPOT may not be useful tools for establishing the severity of pain. Attempts to validate the tools as a method of assessing pain severity have not been successful.41

Summary and Recommendations for Further Research

In summary, internal consistency and interrater reliability in several studies have indicated the reliability of both the CPOT, and to a lesser extent, the BPS. Construct validity of both tools has been indicated in multiple populations of patients. However, the CPOT has been tested in both verbal and nonverbal patients in ICUs, a characteristic that may extend its applicability to practice. In addition, the evidence indicates that the BPS may lack specificity in the detection of pain in noncommunicative critically ill adults, a phenomenon not found in the testing of the CPOT. A comparison of the evidence supporting the use of the BPS with that supporting use of the CPOT suggests that the CPOT may be a more specific, reliable tool in nonverbal critically ill patients.

Although multiple observational pain scales are available, few have been reliable and valid across a multitude of patient populations and settings.
<table>
<thead>
<tr>
<th>Research study</th>
<th>Reliability findings</th>
<th>Validity findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavioral Pain Scale (BPS)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payen et al, 2001</td>
<td>Interrater reliability: 0.94 (Cronbach $\alpha$)</td>
<td>Discriminate validity: Painful procedure scores (4.6-5.2) compared with nonpainful procedure scores (3.3-3.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Construct validity, test-retest procedure: BPS score increased from 3.9-3.3 at rest to 4.0-4.8 during procedure</td>
</tr>
<tr>
<td>Aïssaoui et al, 2005</td>
<td>Internal consistency: 0.72 (Cronbach $\alpha$)</td>
<td>Construct validity: mean score, 3.9 (SD, 1.1) at rest and 6.8 (SD, 1.9) during procedures</td>
</tr>
<tr>
<td></td>
<td>Interrater agreement: 0.95 (ICC)</td>
<td></td>
</tr>
<tr>
<td>Young et al, 2006</td>
<td>Interrater agreement: 82%-91% at rest, 36%-46% during procedure</td>
<td>Construct validity: 3.36 at rest vs 5.02 after painful procedure</td>
</tr>
<tr>
<td>Ahlers et al, 2008</td>
<td>Interrater reliability: 0.67 (95% CI, 0.54-0.80)</td>
<td>Discriminate validity: 3.23 at rest vs 3.38 after nonpainful procedure</td>
</tr>
<tr>
<td>Ahlers et al, 2010</td>
<td>Internal consistency: Cronbach $\alpha = 0.63$-0.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Construct validity: moderate correlation ($r = 0.55$, $P &lt; .001$, $n = 57$) when nurse-assigned NRS score compared with BPS score</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Critical-Care Pain Observation Tool (CPOT)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gélinas et al, 2004</td>
<td>Interrater reliability: weighted $\kappa$ coefficient = 0.52-0.88</td>
<td>Discriminate validity: Painful procedure scores compared with nonpainful procedure scores at rest ($t$-test, $P &lt; .001$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Construct validity: strong correlation of VRS-4 and BPS scores ($r = 0.67$, $P &lt; .001$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discriminate validity: 3.4 at rest vs 5.1 during painful procedure</td>
</tr>
<tr>
<td>Gélinas and Johnson, 2007</td>
<td>Interrater reliability: 0.80-0.93 (ICC)</td>
<td>Positive predictive value of CPOT 85.7% when compared with self-reported FPT values</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marmo and Fowler, 2010</td>
<td>Interrater reliability: 0.89 (Cronbach $\alpha$)</td>
<td>No data</td>
</tr>
<tr>
<td></td>
<td>Interrater agreement: 56%-100%</td>
<td></td>
</tr>
<tr>
<td>Wibbenmeyer et al, 2011</td>
<td>Interrater reliability: 0.63, $P &lt; .001$ (Pearson correlation coefficient)</td>
<td>Discriminate validity: 0.27 at rest vs 0.56 during stimuli</td>
</tr>
<tr>
<td>Vázquez et al, 2011</td>
<td>Interrater reliability: $\kappa = 0.79$</td>
<td>Discriminate validity: mean score at rest, 0.27 (SD, 0.64); mean score during procedure, 1.93 (SD, 1.41)</td>
</tr>
<tr>
<td></td>
<td>Interrater agreement 97%-100%</td>
<td></td>
</tr>
<tr>
<td>Gelinas et al, 2011</td>
<td></td>
<td>No data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nonverbal Pain Scale (NVPS)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Odhner et al, 2003</td>
<td>Internal consistency: Cronbach $\alpha = 0.78$</td>
<td>Construct validity: 0.86, ($P = .05$) when compared with FLACC scores</td>
</tr>
<tr>
<td>Kabes et al, 2009 (revised version)</td>
<td>Interrater reliability: 90%-94.7% (Cronbach $\alpha$)</td>
<td>Discriminate validity: Original NVPS, 135.86; $P &lt; .001$; $n = 122$ Revised NVPS, 145.05; $P &lt; .001$; $n = 122$</td>
</tr>
<tr>
<td>Marmo and Fowler, 2010</td>
<td>Interrater reliability: 0.89 (Cronbach $\alpha$)</td>
<td>No data</td>
</tr>
<tr>
<td></td>
<td>Interrater agreement: 75% on face dimension</td>
<td></td>
</tr>
<tr>
<td>Wibbenmeyer et al, 2011</td>
<td>Internal consistency: Cronbach $\alpha = 0.80$</td>
<td>Discriminate validity: Mean score “at rest,” 0.19 Mean score “noxious stimuli,” 0.44</td>
</tr>
<tr>
<td></td>
<td>Interrater reliability: 0.59 (Pearson correlation coefficient)</td>
<td>Construct validity: correlation of NVPS and NRS poor, 0.38, ($P = .1$)</td>
</tr>
</tbody>
</table>

**Table 5**

Abbreviations: FLACC, Face, Legs, Activity, Cry, Consolability scale; FPT, Faces, Pain, Thermometer scale; ICC, intraclass coefficient; NRS, Numerical Rating Scale; VRS-4, Verbal Rating Scale 4.
Pain assessment in noncommunicative critically ill adults remains a work in progress. By design, observational pain scales require the presence of a spontaneous, neuromuscular-mediated physical response that can be observed by a third party. Therefore, patients who are physically unable to produce such a response, such as patients who have quadriplegia or neuromuscular disorders or are receiving neuromuscular blocking agents, cannot be assessed with these tools. Patients with compromised hemodynamic status, who account for a large proportion of critically ill adults, are also at risk; these patients were often excluded in the aforementioned studies. Therefore, the validity of observational pain scales in patients with hemodynamic instability needs to be further explored.

Few investigations have addressed the assessment and treatment of chronic pain in the ICU. Furthermore, delirium and agitation are common in the ICU; 28% to 87% of ICU patients experience one or both of these conditions at least once during their stay in the unit. The presence of delirium or agitation makes ascertaining a pain level difficult. Currently, no pain assessment tools have been validated in patients who are critically ill or are being treated with mechanical ventilation who are concurrently delirious or agitated.

Finally, pain is a complex, personal experience that cannot easily be adequately described and treated without a patient’s self-report. Pain management providers often use describing words such as “burning” or “hot” to indicate specific types of pain (eg, neuropathic pain). Without the benefit of that type of subjective data, patients may be at risk for receiving a less effective pharmacological agent (eg, opioids for the treatment of neuropathic pain). Future studies on strategies that promote self-report of pain in critically ill patients who are sedated and/or being treated with mechanical ventilation would contribute much to the advancement of pain assessment.

Conclusion

The rates of uncontrolled pain in critically ill patients remain unacceptably high. A systematic assessment of pain should be done routinely, and self-report by the patient should be the primary basis for pain evaluation whenever possible. The routine assessment of pain with an observational pain assessment instrument can decrease ICU length of stay; decrease the duration of mechanical ventilation; and increase the satisfaction of patients, patients’ family members, and health care providers. Of the available observational pain scales, the CPOT has shown superior reliability and validity when used in nonverbal critically ill adults. However, the CPOT should be used cautiously in evaluating patients who have chronic pain or concurrent delirium. More research is indicated in the assessment of pain in patients incapable of spontaneous neuromuscular movement and in those with chronic pain.

Financial Disclosures

None reported.

References


CNE Test  Test ID C1333: Observational Pain Scales in Critically Ill Adults

Learning objectives: 1. Describe the impact of uncontrolled pain on critically ill patients  2. Identify 2 validated pain scales for critically ill patients  3. Describe instrument reliability and validity when considering pain scales for critically ill patients

1. Which of the following statements best describes the pain response?  
a. Inhibition of nociceptors transmits pain impulses to the ventral horn of the spinal cord.  
b. Chronic and acute pain pathophysiology is related to direct tissue injury transmitted by peripheral nerves.  
c. After tissue injury, inflammatory mediators stimulate pain impulses to be sent to the dorsal horn of the spinal cord.  
d. The inhibition of the somatosensory cortex results in the pain localization.

2. Which of the following statements describes the Adult Nonverbal Pain Scale (NVPS)?  
a. Both behavioral and physiologic data demonstrate validity of the NVPS.  
b. The domains of NVPS include behavioral, physiologic, and autonomic scoring between 0 and 10.  
c. The basis of the NVPS scale is well-studied in adults and children.  
d. An NVPS score of 2 in the physiologic category indicates a change in heart rate of <20 beats/min.

3. Which of the following statements most accurately reflects the issue of pain in critically ill patients?  
a. Sedation of critical care patients prevents the pain response from occurring.  
b. Critical illness blunts the pain response in critically ill patients.  
c. Studies indicate no reliable scales for pain assessment in critically ill patients.  
d. Pain tools for patients who receive mechanical ventilation are advocated by multiple professional societies.

4. Which of the following statements best describes the interrater reliability of the NVPS?  
a. The revised version of the NVPS retained the ≥90% interrater reliability.  
b. The internal consistency of the NVPS in several studies was fair; however, interrater reliability remained high.  
c. Intermittent reliability means that the patient self-reported a similar pain rating as the one observed.  
d. Facial expressions in the NVPS demonstrated the highest score of interrater reliability.

5. Which of the following statements describes the correct usage of the Critical-Care Pain Observation Tool (CPOT)?  
a. Pain observation include facial expression, muscle tension, upper limb movements, and reaction to endotracheal suctioning.  
b. This tool is designed to be used with both intubated and nonintubated critical care patients.  
c. The only observational tool that provides a reliable pain intensity assessment.  
d. This tool should be used for patients with delirium only.

6. Which of the following statements most accurately describes the review of pain scales in the literature?  
a. Physiologic data are the most sensitive indicators for pain presence.  
b. There is sufficient research for the NPAT pain scale to indicate its usage.  
c. There is support for the use of the Behavioral Pain Scale (BPS) and CPOT in noncommunicative critically ill patients.  
d. The BPS and CPOT are useful in establishing pain severity in noncommunicative critically ill patients.

7. Which of the following accurately describes study findings regarding the BPS and the CPOT?  
a. Multiple studies support the use of the BPS and CPOT due to reliability and validity of the scales.  
b. Patients with impaired neuromuscular response can be adequately assessed with the BPS and CPOT.  
c. There is a strong correlation between observational pain scales and patient self-reports.  
d. Sensitivity and specificity for both the BPS and CPOT are reported in study findings.

8. Which of the following statements correctly reflects ventilator considerations with observational pain assessment?  
a. Tolerance of mechanical ventilation is described with higher numbers on the observational pain scales.  
b. Coughing is not a consideration in scoring an observational pain scale.  
c. Asynchrony and frequent alaraming is addressed in the CPOT.  
d. Tolerance of endotracheal suctioning is the best observation to determine pain.

9. Which of the following statements most accurately describes uncontrolled pain in critically ill patients?  
a. The physiologic response to pain can be adequately controlled with steroids and sedation.  
b. High rates of uncontrolled pain are common in the intensive care unit.  
c. Psychosocial effects of uncontrolled pain are rarely observed in critically ill patients.  
d. No methods to appropriately address pain except for patient self-reporting currently exist.

10. Which of the following is the most reliable tool in nonverbal critically ill patients?  
a. NVPS  
b. CPOT  
c. BPS  
d. PAIN

11. Which of the following statements is true about observational pain assessment tools?  
a. They are easily incorporated into practice because minimal education is needed.  
b. They are not reliable in assessing the presence of pain in critically ill patients.  
c. They can be used to standardize pain assessment for patients receiving mechanical ventilation.  
d. They are mandated by regulatory agencies.

12. Which of the following statements accurately describes the assessment of body movements in the CPOT?  
a. Absence of movement indicates absence of pain.  
b. Fully bending upper extremities and finger flexion indicates a score of 3.  
c. Muscle tension must be evaluated by the lower extremities.  
d. Restlessness is indicated by movements such as attempting to sit up and pulling tubes.

Test answers: Mark only one box for your answer to each question. You may photocopy this form.

1. a  2. a  3. a  4. a  5. a  6. a  7. a  8. a  9. a  10. a  11. a  12. a

Test ID: C1333  Form expires: June 1, 2016  Contact hours: 1.0  Pharma hours: 0.0  Fee: AACN members, $0; nonmembers, $10  Passing score: 9 correct (75%) Synergy CERP Category A  Test writer: Cheri Blevins, RN, MSN, CCN, CCNS

AMERICAN ASSOCIATION of CRITICAL-CARE NURSES

For faster processing, take this CNE test online at www.ccnonline.org or mail this entire page to: AACN, 101 Columbia, Aliso Viejo, CA 92656.

Program evaluation  
Yes No  
Objective 1 was met  
Objective 2 was met  
Objective 3 was met  
Content was relevant to my nursing practice  
My expectations were met  
This method of CNE is effective for this content  
The level of difficulty of this test was:  
easy  medium  difficult  
To complete this program, it took me ___ hours/minutes.

Name ____________________________ Member #__________________________
Address ____________________________
City, State, ZIP ____________________________
Country ____________________________ Phone ____________________________
E-mail ____________________________
RN Lic. 1/2 ____________________________ RN Lic. 2/2 ____________________________
Payment by:  
Visa  M/C  AMEX  Discover  Check ____________________________
Card #: ____________________________ Expiration Date ____________________________
Signature ____________________________

The American Association of Critical-Care Nurses is accredited as a provider of continuing nursing education by the American Nurses Credentialing Center’s Commission on Accreditation. AACN has been approved as a provider of continuing education in nursing by the State Boards of Nursing of Alabama (#ABNP0062), California (#01036), and Louisiana (#ABR2). AACN programming meets the standards for most other states requiring mandatory continuing education credit for relicensure.