MU Sinclair School of Nursing
University of Missouri

MASTER’S EXAMINATION
Cover Sheet

School of Nursing Assigned ME Number: FS2011-1328

My area of specialization is: Adult Critical Care

The problem, within my area of specialization, which I have addressed, is: Pain assessment in the non-communicative, critically ill adult

I would suggest the faculty readers listed below. The MS exam committee assigns the readings of examinations to faculty on the basis of availability and/or appropriateness of readers. The committee cannot guarantee that your examination will be assigned to faculty you list.

- Dr. Louis Miller
- Dr. Rebecca Johnson
- Dr. Candice Books
- Dr. Amy Vogelsmeier
- Dr. Tina Bloom
Pain Assessment in the Non-Communicative Critically Ill Adult

FS2011-1328

University of Missouri
Abstract

Pain is a common and distressing symptom in critically ill patients. Critically ill patients experiencing high levels of pain are at risk for a plethora of negative psychological and physiological consequences, some of which may be life-threatening. A systematic assessment of pain is difficult in intensive care units due to the high percentage of patients who are non-communicative and unable to self-report pain. Although several tools have been developed to identify objective measures of pain, there are currently no recommendations that identify which assessment tool is most appropriate for this patient population. A comprehensive literature search was completed to identify relevant evidence pertaining to the reliability and validity of available observational pain scales. The evidence was evaluated and synthesized to identify the ideal instrument for assessing pain in the non-communicative critically ill patient. While the Critical Care Pain Observation Tool (CPOT) instrument has demonstrated superiority in reliably identifying pain in this patient population, pain assessment for those individuals incapable of spontaneous neuromuscular movements or in those individuals with concurrent conditions, such as chronic pain or delirium, remains an enigma.
Pain Assessment in the Non-Communicative Critically Ill Adult

Clinical Implications of Pain

Pain is a significant problem in the intensive care unit (ICU), and inadequate pain assessment and management has been linked to increased morbidity and mortality (Shannon & Bucknall, 2003). The physiologic response to pain is almost universally negative, causing potentially fatal hemodynamic instability, alterations in immune system functioning, hyperglycemia, and increased release of catecholamine, cortisol, and antidiuretic hormone secretions (Puntillo et al., 2004). Moreover, uncontrolled pain has been implicated in a variety of psychosocial effects including depression, anxiety, delirium, post traumatic stress disorder, and disorientation (Jacobi et al., 2002). Despite the acknowledgement in the literature that pain is a common stressor in the ICU, high rates of uncontrolled pain remain common in the critically ill (Campbell & Happ, 2010). This can be attributed to the presence of circumstances, such as mechanical ventilation or hemodynamic instability, that preclude the assessment of pain by self-report. Despite the availability of strong evidence that documentation of pain assessment improves pain management and decreases patients’ pain, there has been no universally recommended pain assessment instrument for use in critically ill patients incapable of self-report (Shannon & Bucknall, 2003).

Background and Significance

McCaffery, in 1968, provided the salient definition of pain, that is, “pain is whatever the experiencing person says it is, existing whenever the experiencing person says it does” (McCaffery & Pasero, 1999, p. 24). In 1994, the International Association for the Study of Pain (IASP), refined and expanded this definition, reporting “pain is an unpleasant sensory and
emotional experience in association with actual or potential tissue damage, or described in terms of such damage” (IASP, 1994, p. 209). 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) (Siedlecki, 2009).

Nearly five million people are admitted to the intensive care unit annually (Pronovost & Goeschl, 2005) and an estimated 71% of those patients remember experiencing pain during their stay (Klein, Dumpe, Katz, & Bena, 2010). Pain is one of the most common symptoms present in the critically ill and is experienced by each patient in a unique manner (Puntillo, Smith, Arai, & Stotts, 2008). Critically ill patients are predisposed to experiencing pain both by nature of their pathophysiologic process (Blakely & Page, 2001) and by the high frequency of therapies and procedures that they undergo (Summer & Puntillo, 2001). Painful procedures, such as turning, tracheal suctioning, catheter insertion and sheath removal, are performed commonly in the intensive care unit, and precipitate acute pain (Cade, 2008). In addition, many critically ill patients have a history of chronic pain, which complicates assessment and treatment (Curtiss & Haylock, 2001). Furthermore, when pain is present in the critically ill patient it is more likely to be of moderate to severe intensity and multidimensional in nature (Puntillo et al., 2010).

Pain increases the levels of endogenous catecholamines and stress hormones, resulting in physiologic changes that can compromise hemodynamic stability in the critically ill patient (Klein et al., 2010). Pain-induced reflex responses may alter respiratory mechanics, increase cardiac demands, and cause contraction of skeletal muscles, muscle spasms, and rigidity (Puntillo et al., 2004). Uncontrolled pain has also been linked to the development of post
traumatic stress disorder (PTSD), hypoventilation, increased risk of myocardial ischemia, and delirium (Wilson, Grande, & Hoyt, 2007). Furthermore, pain is one of the most common stressors in the ICU, resulting in a stress response characterized by tachycardia, increased myocardial oxygen consumption, hypercoagulability, immunosuppression, and persistent catabolism (Chanques et al., 2006).

The routine use of an appropriate assessment of pain has been mandated by the Agency for Healthcare Research and Quality (AHRQ) and The Joint Commission (TJC) (Gelinas, Fortier, Viens, Fillion, & Puntillo, 2004). Moreover, nearly every related professional organization, including the American Association of Critical Care Nurses (AACN), the American College of Chest Physicians (ACCP), the Society for Critical Care Medicine (SCCM), and the American Society for Pain Management (ASPM), has advocated for the implementation of standardized pain assessment tools that include behavioral indicators in sedated and ventilated patients incapable of self-report, or in those who self-report may be unreliable (Pun & Dunn, 2007). The implementation of routine assessment of pain in the critically ill has demonstrated increases in patient satisfaction regarding perceived level of pain control (Gelinas et al., 2004). Furthermore, studies have demonstrated that appropriate treatment of pain and anxiety is associated with decreased length of mechanical ventilation and the rate of nosocomial infections (Chanques, et al., 2006). Additionally, systematic pain assessment in the critically ill has been shown to decrease ICU length of stay (Payen, Bosson, Chauses, Mantz, & Labarere, 2009).

The first step in providing adequate pain relief for patients is the systematic and consistent assessment and documentation of pain (Arif-Rahu & Grap, 2010). However, many patients in the intensive care units are unable to self-report pain due to the presence of mechanical ventilation, sedation or critical illness. Even in patients capable of reporting their
pain, the use of self report as a method of pain assessment is fraught with limitations (Schiavenato & Craig, 2010). Although patient self-report is still considered the *gold standard* of pain assessment, the distress associated with pain can generate a broad range of observable behaviors that can indicate the presence of pain, such as facial expression, physical movement, or autonomic responses (Cade, 2008). Many times, these behaviors are used by clinicians to signify the need for analgesia. Yet, when nurses are used as proxy reporters, the correlation between the nurses’ assessment of pain and the patient’s pain rating by self-report is low (Puntillo et al., 2008). In addition, family members do poorly at identifying pain in the patient as well, reporting the presence of pain only 53% of the time (Payen et al., 2001).

In recent years, with the focus placed by regulatory agencies on the identification and treatment of pain, there has been an impetus toward the development of behavior-based pain scales to assess pain in the patient incapable of self-report (Pasero & McCaffery, 2005). The tools are based on the identification of behaviors, such as facial expressions, vocalizations, withdrawal reflexes and other motor movements, which are associated with the existence of pain (Cade, 2008). Several behavior-based instruments have been developed, including the Adult Nonverbal Pain Scale (NVPS), the Behavioral Pain Scale (BPS), the Critical-Care Pain Observation Tool (CPOT), and the Pain Assessment and Intervention Notation (P.A.I.N.) algorithm. Despite the many tools available to identify pain in the nonverbal patient, the evidence comparing these tools is sparse.

**Purpose and Organization**

The purpose of this paper is to identify the current tools available for assessing pain in the critically ill patient population and to analyze the data supporting the reliability and validity of
these instruments. Usefulness in practice and feasibility will also be explored. Each instrument is described, and a subsequent evaluation of the available relevant evidence is provided based on study design and methods, study sample, theoretical foundation (if any), major findings, and limitations of the literature. Following the findings of this analysis, a comprehensive synthesis of the evidence is described. Next, gaps in the literature are identified and discussed. Finally, implications for future research aimed at improving the understanding of pain assessment in the non-communicative critically ill are described. If a valid, reliable, easy to use instrument could be identified, the effects of pain on patient and family satisfaction, hemodynamic stability and patient outcomes may be mitigated.

Search Strategy

To address this issue, a literature search was conducted utilizing OVID, PubMed, Medline and CINAHL databases. The key words pain, assessment, tool, instrument, scale, intensive care, critical care, and critically ill were used. Articles must have been published within the last 10 years, and available in the English language to be considered for inclusion in the review. Articles were limited to those from professional, scholarly, peer reviewed nursing and health care journals. Articles containing a pediatric focus, addressing pain assessment outside of the intensive care unit, or focused on the evaluation of chronic pain were excluded. A total of 1,120 articles were found, with 17 articles meeting inclusion criteria.

Review of the Literature

A multitude of tools have been developed to recognize objective behaviors that indicate pain in the non-communicative patient. Objective pain measures, also called behavioral-based pain scales, are used by the observer (usually a healthcare provider, most commonly a registered
To identify behaviors that are associated with the presence of pain. In most tools, a higher score indicates a higher probability that the patient is experiencing pain. These tools have been designed and tested in the critically ill population, pediatrics, and in those with dementia or other diagnoses that affect the patient’s ability to provide self-report of pain. This review will focus on those tools that have been developed for and tested in the adult critical care patient population. These instruments include the Pain Assessment and Intervention Notation (P.A.I.N.) 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 and Physiologic Indicators (CPOT).

**The Pain Assessment and Intervention Notation (P.A.I.N.) Algorithm**

The Pain Assessment and Intervention Notation (P.A.I.N.) algorithm was developed in 2001 and consists of three parts: pain assessment, the ability of the patient to tolerate opioids and guidelines for analgesic treatment decisions and documentation (Puntillo, Stannard, Miakowski, Kehrle, & Gleeson, 2002). The pain assessment section of the tool includes dimensions that assess for the presence or absence of pain related behaviors (movement, facial cues, posturing) as well as physiological indicators (increased heart rate, respiratory rate, and blood pressure, perspiration or pallor). The observer then provides their subjective assessment of the patient’s pain intensity on a 0-10 scale, where 0=no pain and 10=worst pain imaginable. This was compared to the patient’s self-report of pain intensity on the same scale. The RN participants were also asked to complete an evaluation of the tool that assessed the helpfulness of the tool in assessment of pain, the usefulness of the tool in clinical practice, the participant’s perceived accuracy of the tool, and its perceived effect on practice.
The initial testing of the tool included a study sample of 11 RNs who used the instrument to assess 31 postoperative patients in the ICU or post anesthesia care unit (PACU). In evaluating the data from this study, the researchers found that the majority of the participants felt the tool was helpful in providing a “systematic approach” to pain assessment and guidance of analgesic administration (Puntillo et al., 2002, p. 311). However, four of the 11 participants felt the P.A.I.N. tool was not helpful, believing it was “too long and cumbersome” to be used in “everyday practice” (Puntillo et al., 2002, p. 311). Reliability and validity of the tool were not explored. While the researchers found that the instrument could be a useful training modality for beginning ICU nurses, the length of the tool has limited its clinical utility and no further testing has been done (Li, Puntillo, & Miakowski, 2008).

**The Nonverbal Pain Assessment Tool (NPAT)**

The Nonverbal Pain Assessment Tool (NPAT) was developed for use in the nonverbal adult patient in the ICU and is composed of five observational domains including: emotion, movement, verbal cues, facial cues, and positioning/guarding (Klein et al., 2010). Two separate scoring systems are provided on the instrument, and the observer selects the appropriate scale based on the patient’s ability to produce vocal cues. Scoring ranges from 0-10 points with higher numbers indicating higher severity of pain. The initial testing of the NPAT consisted of three phases of validity testing and two revisions of the tool. The third phase of validity testing analyzed the correlation between the nurse assigned NPAT score and the patient’s self-reported pain score in a convenience sample of 50 general medical/surgical, postoperative patients. The data was collected by two independent observers, one of whom collected the patient’s self-reported pain score on the numerical rating scale (NRS), and another, who assigned the patient’s NPAT score.
The study authors reported a high agree of interrater reliability (K=.72, 95% CI) in the testing of the phase three NPAT (Klein et al., 2010). In addition, anecdotal evidence suggests that the NPAT is consistent and easy for nursing staff to administer. Criterion validity was moderately strong when the NPAT was compared to the gold standard of self-report (.66, 95% CI). Although designed for non-communicative critically ill patients, the NPAT has not been validated in this population, as the initial testing of the third phase NPAT was conducted on verbal, medical surgical patients. While the first two phases of reliability and validity testing of this instrument were conducted in intensive care units, the validity of the NPAT was found to be weak (.31, 95% CI and .21, 95% CI, respectively). Moreover, the study sample was small, consisting of 30 postoperative patients and a single center design was used. No further reliability or validity testing of the NPAT has been completed.

**The Adult Nonverbal Pain Scale (NVPS)**

The Adult Nonverbal Pain Scale (NVPS) was developed in 2002, based on a previously designed observational pain scale called the Faces, Legs, Activity, Cry, Consolability (FLACC) scale validated in the pediatric patient population. Like the P.A.I.N. and NPAT instruments, the NVPS contains the behavioral dimensions of facial expression, activity and guarding. In addition, the NVPS also contains physiologic dimensions that are designed to identify pain-related changes in heart rate, blood pressure and respiratory rate, which are graded in severity. For example, a systolic blood pressure increase of greater than 20mm Hg from baseline would be scored a “1”, whereas a systolic blood pressure increase of greater than 30mmHg from baseline would receive a “2”. The final domain of the NVPS includes additional physiologic autonomic indicators like 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 administered. The authors concluded that the NVPS was a valid observational pain scale in this patient population based on the correlation of the NVPS to the FLACC instrument which was found to be high (.86, p=0.05). However, the use of the FLACC as a *gold standard* measurement in this study cannot be supported, as the FLACC was developed for use in pediatric patients and has not been validated in adults. The NVPS authors also reported high internal consistency (\(a=0.78\)), as well as high interrater reliability, although no statistical analyses were reported (Odhner, Wegman, Freeland, Steinmetz & Ingersoll, 2003). In addition, the non-experimental design and the relative homogeneity of the study sample threaten the validity and generalizability of this study.

In another study by Kabes, Graves, and Norris (2009), a revised version of the NVPS was compared against the original NVPS for reliability and validity testing. The revised NVPS, as reported by Wegman (2005), includes a new “respiratory” category that replaces the physiologic II dimension on the original scale. The respiratory 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. In this non-experimental study design, Kabes et al. (2009) utilized registered nurses in a trauma/surgical intensive care unit to assess patients before, during, and at rest after, a painful nursing procedure. Both the new respiratory dimension and the original physiologic II dimension were assessed concurrently with the remainder of the instrument during 121 paired observations.

The study authors reported a high level of inter-rater reliability in nearly all comparisons (90% or more in 94.7% with original NVPS and 90.8% with the revised NVPS). The original scale demonstrated internal consistency levels, as indicated by Cronbach’s alpha values, of 0.36
(prior), 0.62 (during), and 0.71 (after). The revised scale resulted in Cronbach’s alpha values of 0.36 (prior), 0.72 (during) and 0.71 (after). The authors found that all items on the scale, with the exception of the physiology II dimension, showed moderate correlations. 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=122; revised, 145.05, p<.001, n=121) (Kabes et al., 2009). These findings support the construct validity of both instruments, but also require that the data collectors know which stage the patient being observed is in at the time of collection. Knowing the stage (before, during, after) of the painful procedure may have influenced the observers rankings. In addition, correlations between NVPS scores and self-report scores were not studied. The authors concluded the revised version of the NVPS demonstrated validity in the assessment of pain in nonverbal critically ill adults in this study.

In a repeated-measures study design, Marmo & Fowler (2010) sought to validate the NVPS in the critically ill post-open heart surgery population. Utilizing a convenience sample of 25 patients, the NVPS, CPOT, and FLACC were administered for a total of 300 paired observations (completed by two different data collectors concurrently). A total of 280 observations were needed for a medium (0.30) effect size at a .05 level of significance and 80% power. The instruments were completed before, during and after a painful procedure (endotracheal suctioning, turning), similar to the Kabes et al. (2009) study. The data collectors were trained to administer all tools prior to the study onset and 85% agreement between investigators was obtained prior to obtaining study data.

The study authors found that the NVPS was very reliable, with a Cronbach alpha coefficient of 0.89; however, there was a high occurrence of disagreement between nurse raters on the face dimension of the NVPS (25% of the total observations). The authors concluded that
the “CPOT appears to be a better tool to detect pain in post-open heart surgery adults”, however, this conclusion cannot be supported with the data reported (Marmo & Fowler, 2009, p. 139). While the authors acquired data before, during, and after a painful event, there is no data provided comparing these scores on the respective instruments. These correlational findings are necessary to establish the validity of the tools. Therefore, the study findings only suggest that the tool is reliable and internally consistent, but do not establish validity in pain identification in this population.

However, a recent study by Wibbenmeyer et al. (2011) did investigate both the validity and reliability of the NVPS in a convenience sample of 38 burn patients. A total of 225 paired assessments were completed by nursing staff that were “briefly educated” (Wibbenmeyer et al., p. 53) on the appropriate use of both the CPOT and NVPS scales. The assessments were completed while the patient was at rest and then again shortly after a painful stimulus (wound care, physical/occupational therapy) was applied. Because all of the patients in this study were able to verbally respond, a self-report of pain intensity was also obtained after administration of the aforementioned tools.

The investigators found that the NVPS demonstrated good internal consistency, with a Cronbach’s alpha score of 0.80. Inter-rater reliability was found to be fair at 0.59 using the Pearson correlation coefficient, however, this may have been due to the limited education the data collectors received on use of the instrument prior to the study onset. The NVPS established discriminate validity with the mean at rest scores being 0.19 and increasing to 0.44 after the noxious stimulus was applied. Finally, the authors reported that the NVPS was a poor predictor of pain intensity, with the correlation of NVPS score and numerical rating scale (NRS) being poor (0.38, p<.01). This study demonstrated that the NVPS is a good tool for identifying pain in
the burn population, but not discerning pain intensity. The poor inter-rater reliability shown in this study may likely be due to the limited amount of training provided on instrument use, as no data was collected on inter-rater agreement prior to the data collection period. Finally, in order to collect self-report data for comparative purposes, the patients in the sample must have been able to verbalize pain. The NVPS is designed for patients who are nonverbal, so the results of this study may not be fully applicable to nonverbal patient populations.

The previous studies sought to examine the reliability and validity of the NVPS in various patient populations. In a study by Topolovec-Vranic et al. (2010), the investigators examined the effect of implementing the NVPS on patient satisfaction and nursing confidence in pain assessment in the nonverbal patient. Utilizing the staff and patients in a neurosurgical/trauma ICU, the study investigators employed two questionnaires and chart reviews to obtain outcome data at baseline, during implementation of the NVPS, and four weeks after implementation. The patients present on the unit during these time periods were asked to complete a 12 item questionnaire that assessed the patient’s perceptions of their pain management during their stay in the ICU. An additional 10-item staff questionnaire was administered to staff to assess their supposed comfort level in assessing pain, the perceived barriers to implementing a new pain scale in the ICU, and their opinions about the NVPS (ease of use, barriers to use, usefulness of the tool).

A total of 52 patients had recollection of their ICU stay and completed questionnaires. No statistically significant differences were found in the “worst pain during your ICU stay” scores or the intensity of pain during the ICU stay. Although a trend toward a decreased levels of severe pain post NVPS implementation was noted (55% vs 35% patients reporting severe pain). The time it took for the patient to receive pain medication also showed a positive trend
Increased frequency of pain assessment documentation was noted after implementation (29% vs 59%, p<.001) although this may have been due to an increased awareness of the need to assess pain following the NVPS education sessions that were provided to the staff prior to the data collection period. 78% of the staff reported that the NVPS was “easy” or “very easy” to use, and the improvement in the nurses’ perceived confidence in assessing pain in sedated or nonverbal patients was significant (57% pre implementation vs 81% post implementation, p<.02) (Topolovec-Vranic et al., 2010). While this study demonstrated an increase in nursing confidence and frequency of pain assessment documentation, both of these items could be attributed to the initial educational session and not to the NVPS specifically. Moreover, no significant improvements were shown in patient satisfaction after the implementation of the NVPS. Therefore, the effects of implementing the NVPS on patient and staff satisfaction require further investigation.

The Behavioral Pain Scale (BPS)

The Behavioral Pain Scale (BPS) was developed in 2001 based on the work of Puntillo et al. (1997) that identified several unique behaviors present in patients after undergoing a noxious stimulus. The BPS is composed of three 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 in a convenience sample of 30 mechanically ventilated patients utilizing a quasi-experimental design (Payen et al., 2001). The BPS scores were assessed at rest and then during a painful procedure (defined as application of compression stockings or a central venous line dressing change) in group 1, or during a nociceptive procedure (endotracheal suctioning or
turning) in group 2. The data collectors were also asked to complete a survey related to ease of use and time required to complete the assessment.

Assessments completed at rest had a high percentage of no response (score of 3) without a significant difference in group 1 or group 2. However, painful procedures resulted in much higher BPS scores (4.6-5.2) than in the non-painful procedure group (3.3-3.7). A test-retest procedure was then performed in 31 cases, which resulted in BPS scores that increased from 3.0-3.3 at rest to 4.0-4.8 during a painful procedure, thereby establishing construct validity (Payen et al., 2001). Inter-rater agreement was reported as a Cronbach’s alpha value of 0.94. A majority (24 of the 28) of the evaluators reported the BPS was easy to use, and took 2-5 minutes to complete on average. A small portion of the evaluators expressed concerns that the tool was somewhat complex to complete. In addition, the single center design of this design limits the ability to extrapolate these findings to other patient populations and facilities. However, unlike the studies used to validate other observational scales, this study tested the instrument in the population for which it was intended (nonverbal critically ill patients).

In 2005, Aissaoui, Zeggwagh, Zikraoui, Abidi, and Abouqal, investigated pain in 30 general ICU patients utilizing the BPS in a total of 360 observations. All patients included in the study were sedated and mechanically ventilated. In paired, but independent observations, two observers obtained BPS scores at rest and during painful procedures (endotracheal suctioning or peripheral vein cannulation). The BPS was found to have good internal consistency (Cronbach alpha=0.72) and inter-rater agreement was high (0.95) when intraclass correlation coefficient was used for evaluation. Validity was established by the change in BPS scores with averages of 3.9 +/- 1.1 at rest and 6.8 +/- 1.9 during procedures. The elevated BPS scores at rest are interesting and likely support hypotheses provided in previous research that in critically ill
patients may have moderate levels of pain even at rest. The study authors also examined the responsiveness of the BPS, which they reported to be excellent, with an effect size ranging from 2.2 to 3.4.

Another study in 2005 by Young, Siffleet, Nikoletti, and Shaw utilized a descriptive, repeated measures design to analyze the reliability and validity of the BPS. 44 hemodynamically stable, mechanically ventilated patients were assessed before and after a non-painful (eye care) and a painful (repositioning) procedure. The investigators found that the mean BPS score showed a significant increase after the repositioning procedure (3.36 pre-procedure vs 5.02 post procedure. In contrast, there was no significant change in mean BPS score following the eye care procedure (3.23 pre-procedure vs 3.38 post procedure), supporting the discriminate validity of the BPS. However, inter-rater agreement varied widely, with good agreement (82-91%) during rest states, but a marked decrease in agreement following a painful procedure (36%-46%). These study findings are limited by a small, heterogeneous sample, and the exclusion of those patients with hemodynamic instability, of which comprise a large portion of the population in any given intensive care unit.

In an attempt to recreate the findings of Payen et al. (2001) and Aissaoui et al. (2005), Ahlers et al. (2008) obtained 371 independent observations of the BPS and NRS (when possible) in 113 critically ill patients. The prospective observational study utilized a convenience sample of medical surgical ICU patients with assessments completed by two independent data collectors during non-painful procedures. Prior to the collection of data, the data collectors attended a four hour training course in conducting BPS assessments. The BPS was completed on all ventilated patients while the NRS was either scored by the bedside nurse or by the patient if they were able.
The use of a bedside nurse to obtain a NRS score is a questionable practice, as this study demonstrated that the NRS score by the patient and by the nurse correlated only 73% of the time.

Despite the problematic study design, the inter-rater reliability of the BPS was found to be good (0.67, CI 0.54-0.80). The correlation between the NRS nurse and the BPS researcher was moderate (r=0.55, p<.001, n=57). However, it is interesting to note that only 5% of the observations provided an NRS of 0, while 68% of the BPS observations were scored as a 3, indicating no pain. This data illuminates an interesting question in the Payen et al. (2001) study which showed a high proportion of low BPS scores at rest. It is possible that a BPS score of 3 is not truly representative of a pain-free state as previously postulated. Finally, the BPS equates lack of body movement with a pain-free state. This conflicts with the research done by Puntillo et al. (1997) that demonstrated a correlation with decreased movement and increased pain levels.

In a follow up study, Ahlers, van der Veen, van Dijik, Tibboel, and Knibbe (2010) studied the reliability and validity of the BPS in both deeply sedated patients, and patients undergoing conscious sedation for a painful procedure. Using a convenience sample of medical surgical ICU patients, nurses conducted 175 observations in 80 critically ill adults at 4 points: at rest, during a nonpainful procedure, during a painful procedure, and following the procedure. Patients were classified as sedated (not able to communicate during an assessment period) or conscious sedated (able to communicate during the assessment period). The assessments were conducted by two nurses independently and communicative patients were asked to provide a report of their pain intensity via the use of the Verbal Graphic Scale 4 (VRS-4) if able to do so after the BPS scoring was completed.
Inter-rater reliability of the BPS was excellent with Kappa values ranging from 0.80 (sedated patients) to 0.83 (conscious sedated patients), and internal consistency was moderate with Cronbach alpha scores of 0.63 in sedated patients and 0.66 in conscious sedated patients. BPS scores were higher during painful procedures than at rest in both sedated patients (5.1 [95% CI] vs 3.4 [95% CI]) and conscious sedated patients (5.4 [95% CI] vs 3.8 [95% CI]), thus establishing construct validity. The BPS and the VRS-4 demonstrated 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 patient populations. Although the ideal study design would blind the data collectors to the nature of the procedure, limiting bias, this study is extremely well designed in the remaining aspects.

The Critical-Care Pain Observation Tool (CPOT)

The CPOT was developed based on retrospective chart reviews that examined common pain notations and focus groups with ICU clinicians (Li et al., 2008) and is designed for use in both intubated and non-intubated critical care patients. Four behavioral domains, including facial expressions, movements, muscle tension, and ventilator compliance, are scored from 0 to 2, with a total score ranging from 0 (no pain) to 8 (most pain). The CPOT contains operationally defined descriptors and the content validity of all indicators was found to be between 0.88 to 1.0 in analyzing the results of a questionnaire provided to physicians and critical care nurses (Gelinas, Fillion, & Puntillo, 2008). The CPOT was originally developed in French, and the initial reliability and validity testing was conducted in a convenience sample of 105 cardiac surgery patients.
Using a repeated measures design, Gelinas et al. (2006) used trained data collectors to assess the CPOT score of the patient during three periods: at rest, immediately after repositioning (nociceptive procedure), and at recovery (20 minutes after repositioning). Following the completion of the assessment by two observers, the patient was asked to indicate the presence or absence of pain by nodding their head “yes” or “no”. Patients who were found to have delirium (as identified by the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)) were excluded from the study. Inter-rater reliability between the principle investigator and the critical care nurse observer was found to be moderate to high (weighted K coefficient=0.52-0.88) in all assessments. However, this reporting is of limited value as the same two observers were used in all assessments.

Criterion validity was established by the comparison of the patient’s self-report of pain and the concurrent CPOT score. In patients reporting pain, the mean CPOT score ranged from (1.62 to 3.65) and in patients reporting no pain, the mean CPOT score ranged from (0.49-2.11). While the mean CPOT scores were lower in the patients reporting no pain, the CPOT appeared to overestimate pain in a percentage of cases, especially in the immediate post nociceptive procedure period (mean 2.11, SD .90). Discriminate validity was supported by the finding that the CPOT scores were higher during positioning than at rest (t= -9.01 to -15.96, p<.001). While the study sample in this design was large, inter-rater reliability and the homogeneity of the sample were acknowledged limitations.

In a post hoc data analysis of the previous study sample, the developers of the CPOT evaluated the sensitivity and specificity of the CPOT. They found that the CPOT had a sensitivity of 86% and a specificity of 78% during painful procedures. However, the sensitivity of the tool decreased to 83% prior to a painful procedure and 63% following the painful
procedure. Specificity remained high at 83% and 97%, respectively, during these study periods (Gelinas, Harel, Fillion, Puntillo, & Johnson, 2009). While this data supports the authors’ conclusion that the CPOT is most useful for detecting pain during nociceptive procedures, the clinical utility of this data is questionable, as the most valuable tool can identify pain when the clinician caring for the patient may not anticipate it.

The first English version of the CPOT was developed and tested in a group of 30 conscious and 25 unconscious patients in a general ICU employing a cross-over observational design (Gelinas & Johnston, 2007). Additional data was collected in this study, including physiologic data, and the patients self-report of pain utilizing the Faces Pain Thermometer (FPT). Inter-rater reliability was supported in this study with intraclass correlation coefficients ranging from (.80 to 0.93) amongst the 51 registered nurse data collectors. The authors report that discriminate validity was established in this study due to the presence of heart rate (HR) and blood pressure (BP) elevations that increased during turning in accordance with elevated CPOT scores, however, physical motion is known to increase HR and BP to compensate for increased oxygen demand, making this method of establishing validity problematic. However, when the patient’s self reported FPT values where compared with the observer derived CPOT scores, the positive predictive value of the CPOT was found to be high at 85.7%.

The registered nurse participants from the previously described study were also asked to complete an evaluation form addressing the feasibility and clinical utility of the CPOT in practice (Gelinas, 2010). The nurse’s demographic data varied widely, with age ranges from 24-53 years of age and years of nursing experience ranging from 6 months to 28 years. Most nurses were employed on a full-time basis and possessed either a college diploma or a bachelor’s degree in nursing. 100% of the nurse respondents reported that the CPOT directives were clear and that
the tool was easy to use. 78% reported that the CPOT was quick to use, and 72.7% reported that they would recommend the use of the CPOT routinely in practice (Gelinas, 2010).

In the aforementioned study by Marmo and Fowler (2010), the CPOT was again tested in the post-open heart surgery population and was found to have high reliability with a Cronbach alpha score of 0.89. Marmo and Fowler (2010) were also the first to report the internal consistency of the CPOT, which ranged from 56%-100% agreement amongst nurse raters. When compared to the NVPS, the CPOT was found to have less instances of disagreement amongst nurse raters (14% vs 10%). The CPOT was also included in the previously discussed study of pain assessment in burn patients by Wibbenmeyer et al. (2011), who reported a high internal consistency of 0.71 and good discriminate validity (mean scale scores=0.27 at rest to 0.56 after noxious stimulation). As with the AVPS, Wibbenmeyer et al. (2011) reported that the inter-rater reliability of the CPOT was poor (Pearson correlation coefficient of 0.63, P<.0001). However, this could again be due to the limited amount of training the assessors receiving prior to data collection.

In 2011, Vazquez et al. conducted a prospective, repeated measures study in a 12-bed general intensive care unit in Spain. A total of 330 paired observations were completed in a study sample of 98 critically ill patients. Observations were again conducted before, during, and after a repositioning procedure. The study found the inter-rater reliability of the CPOT to be excellent (k=0.79-1), with agreement percentages ranging from 97% to 100% agreement. The CPOT was found to have good discriminate validity, with average at rest scores being 0.27 (SD=0.64) and average during procedure scores being 1.93 (SD=1.41). The researchers also found a trend toward higher CPOT scores in surgical patients (mean-2.02) than medical patients (mean-1.80) although the results were not statistically significant.
Finally, Gelinas et al (2011) completed a pre and post evaluation of the effects of implementing the CPOT in a general intensive care unit in Canada. Utilizing a before and after study design, the researchers examined the inter-rater reliability of the CPOT, as well as the pain assessment, and pain management practices of the nursing staff. Nurses were asked to assess the CPOT score of a videotaped patient prior to implementing the CPOT, and again at 12 months after implementation to assess inter-rater reliability. In the pre-implementation period, inter-rater agreement when scoring patients at rest was high (95-100%), with agreement during turning found to be acceptable (73-91%). In the post implementation period, inter-rater agreement improved to 86-100%. When analyzing pain assessment practices using descriptive statistics, reports of pain assessments were three to four times more frequent in the post implementation period than in the pre-implementation period. Interestingly, the implementation of the CPOT was associated with decreased frequency of sedative and analgesic administrations in the post implementation period. The authors provide two possible explanations for this phenomenon including the increased ability of nursing staff to discern pain from other symptoms (such as anxiety), or the decreased number of trauma patients in the post implementation group due to a change in the center’s trauma designation.

Theoretical Foundations of the Literature

Despite the multitude of literature focused on the validation of observational pain scales in critical care environments, there is a paucity of studies describing a theoretical framework on which the work is based. In fact, of the studies included in the aforementioned literature review, none include a reference to a theoretical perspective as a basis for the study design, or as a means to explain the study conclusion. However, several of these published studies have yielded findings that are consistent with previously described theoretical viewpoints, namely, the Theory
of Unpleasant Symptoms and the Symptom Management Model. Both of these middle range theories are frequently used to describe the management of symptoms, including pain, in the critically ill.

The Theory of Unpleasant Symptoms proposes that the factors associated with a symptom may influence a number of different concurrent symptoms. Therefore, the use of therapeutic interventions aimed at lessening the severity of one symptom may be effective in alleviating other related symptoms. The Theory of Unpleasant Symptoms includes three reciprocal components including the symptoms that the individual is experiencing, the influencing factors giving rise to or affecting the nature of the symptom experience, and the consequences of the symptom experience (Lenz, Pugh, Milligan, Gift, & Suppe, 1997). Symptoms are described in dimensions of intensity, timing, level of perceived distress, and quality. These dimensions are influenced by 3 variables; that of physiologic, psychologic, and situational factors.

The variables described in the Theory of Unpleasant Symptoms provide insight into the pain characteristics in the critically ill. This population frequently presents with alterations in all three of the described factors. By nature of their admission to the intensive care unit, critically ill patients have physiologic alterations that may contribute to their symptom experience. Moreover, there is an overabundance of evidence describing the negative psychological affects that are associated with critical illness including delirium, agitation, post traumatic stress disorder (Wilson, Grande, & Hoyt, 2007). Finally, a host of environmental alterations (such as bright lights, alarms, invasive equipment) and inability to communicate are commonplace in the intensive care unit. The accumulation of these alterations, and the high frequency of concurrent symptoms in the critically ill (such as nausea, vomiting, and thirst) may explicate why a large
The majority of critically patients consistently rank their pain as moderate to severe (Puntillo et al., 2010). The well documented association between the rates of uncontrolled pain and the rates of delirium in the critically ill is also supported by the constructs of the Theory of Unpleasant Symptoms.

The Symptom Management Model contends that in order to effectively manage symptoms, one must have a comprehensive understanding of the symptom experience, symptom management strategies, and symptom outcomes (Sousa, McDermott, & Weiss, 2011). Echoing the Theory of Unpleasant Symptoms, the Symptom Management Model asserts that the characteristics of the person, the health/illness state, and the environment all contribute to the symptom experience, management, and outcome. This contention is supported by the complex nature of pain quality and pain management in the intensive care unit. Several assumptions included in the Symptom Management Model are consistent with literature findings, including that the gold standard of study is based on the patient’s perception of pain severity, and that the nonverbal patient experiences symptoms which are interpreted by the caregiver to assess symptom experience (Dodd et al., 2001). Despite the dearth of theoretical references in the observational pain scale literature, it is clear that many of the aforementioned study findings support the concepts described in these two theoretical frameworks.

Analysis of the Literature

It is apparent from the evidence that more research aimed at finding the ideal instrument to assess pain in the critically ill nonverbal adult is needed. While multiple tools are available, few have been found to be reliable and valid across a multitude of patient populations and settings. However, it is quite clear that some tools have been tested much more than others. The
P.A.I.N. algorithm and NPAT have each been used in only one study and both studies were of limited value. The P.A.I.N. algorithm was found to be too lengthy to be used in clinical practice and no further modifications of the scale have been done. The final version of the NPAT was tested in a group of verbal general medical surgical patients, which grossly limits its generalizability and applicability to non-communicative critically ill patient populations. More appropriate study sampling is indicated before the NPAT can be considered an appropriate tool for pain assessment in the intensive care units.

The original testing of the NVPS is of limited value due to its non-experimental design and the use of the FLACC scale as the gold standard for comparison. While the NVPS is the only tool that includes physiologic data dimensions, these indicators have been found to be some of the least sensitive markers for the presence of pain (Pasero & McCaffery, 2000). In addition, the NVPS developers fail to provide literature supporting the criterion for the physiologic category intensities. Moreover, in the follow up studies by Marmo and Fowler (2010), and Wibbenmeyer et al. (2011), comparing the NVPS to the CPOT, the NVPS was found to have more rater disagreement. However, in both studies, the NVPS was found to have acceptable discriminate validity. It appears that while the NVPS is a valid tool to identify pain, there may be a more reliable tool available for this purpose.

An overwhelming majority of the studies provide support for the reliability and validity of both the BPS and the CPOT in identifying pain in the non-communicative critically ill adult. In regards to reliability, both the BPS and CPOT have been assessed for both internal consistency and inter-rater reliability in more than one well designed study. Both tools have demonstrated good to high reliability in all analyses. However, the CPOT has been found to be more internally consistent than the BPS (Cronbach $a=0.71$-$0.89$ vs. $a=0.63$-$0.66$, respectively)
when study findings are compared. Whether this data translates into a true implication for practice remains to be seen, as very high levels of internal consistency may simply indicate redundancy within the tool (Boyle, 1994).

In addition, both the BPS and CPOT have demonstrated fair to good inter-rater reliability in most studies, with the exception of the Wibbenmeyer et al. (2010) study. This finding may be attributed to the minimal training provided to the raters prior to performing assessment with the CPOT. While the results of inter-rater reliability testing of the tools are comparable, the inter-rater reliability of the CPOT has been more extensively tested (6 studies reporting data, compared with 4 studies reporting data for the BPS). Moreover, the inter-rater reliability data reported for the CPOT is more likely to be accurate, owing to a broader range of statistical analyses employed in the data analysis (ICC, Pearson correlation coefficients, kappa values, percentage of agreement) and more rigorous study designs that included a larger number of raters in the analyses.

Validity analyses conducted in a variety of patient populations exploring construct, criterion, and discriminate validity have been predominately favorable of both the BPS and CPOT, with both tools able to demonstrate an increase in score after a painful procedure was performed. However, the Payen et al. (2001) and Young et al. (2006) studies reported an increase in the BPS score after a non-nociceptive procedure as well, making the specificity of the instrument a concern. Moreover, while the sensitivity and specificity of the CPOT has been investigated (Gelinas et al., 2009), no data regarding sensitivity or specificity of the BPS has been reported. This issue is complicated further by the average “at rest” scores, which were elevated in both tools (BPS 3.7, CPOT 0.27). These findings suggest baseline pain is present in
a majority of the critically ill, which is consistent with a previous study describing characteristics of pain in the ICU (Puntillo et al., 2004).

An important finding of the validity testing of both tools is a lack of strong correlation of the observational pain tools with the patient’s self-report. Only a moderate correlation was found between both instruments and the patient’s rating on self-report scales. The correlation was especially poor when the patient was scored on the observation tool as having “no pain” (BPS score of 3, CPOT score of 0), as very few patients reported having no pain on self-report scales. These findings suggest that while scores increased during nociceptive procedures indicating the presence of increased pain, they may not be useful tools for establishing the severity of pain. Attempts to validate the tools as a method of assessing pain severity (as in the Wibbenmeyer et al. (2010) study) result in poor validity. Moreover, in all studies, by nature of the study design, the assessors were aware of the nature of the procedure (i.e. painful vs. non-painful), which may have introduced some bias into the assigned score.

**Summary and Recommendations for Further Research**

In summary, both the CPOT, and to a lesser extent, the BPS, have demonstrated reliability by internal consistency and inter-rater reliability in several studies. Construct validity of both tools has been demonstrated in multiple patient populations. However, the CPOT has been tested in both verbal and nonverbal patients in intensive care units, extending its applicability to practice. In addition, the evidence indicates that the BPS may lack specificity in the recognition of pain in the non-communicative critically ill adult. The evidence demonstrating the reliability and validity of the CPOT when used in the non-communicative critically ill incapable of self report is of high quality. When compared with the evidence
supporting the use of the BPS, it is clear that the CPOT is a superior tool in this patient population.

Pain assessment in the non-communicative critically ill adult remains a work in progress. While the observational pain assessment tools available provide some insight into the presence of pain, there are still many unanswered questions. 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, as in quadriplegia, neuromuscular disorders, and those receiving neuromuscular blocking agents, are unable to be assessed with these tools. This creates a situation where the patient is at risk for either under treatment of pain or analgesic agent overdose. Patients with hemodynamic compromise are also at risk, as these patients were frequently excluded in the aforementioned studies. These patient groups comprise a large portion of the patient population in many intensive care units.

In addition, there has been little research done into the area of assessment and treatment of chronic pain in the intensive care unit. While several instruments are available to assess pain in the chronic pain patient, none of these tools are designed to assess chronic pain a non-communicative patient who is critically ill (Hamill-Ruth & Marohn, 1999). Furthermore, delirium and agitation are common in the intensive care unit with 28-87% of patients experiencing one or both of these conditions at least once during their stay (Puntillo et al., 2010). The presence of delirium or agitation complicates the issue, making it difficult to discern pain symptoms from that of delirium or agitation. Presently, there are no pain assessment tools that have been validated in the presence of concurrent delirium or agitation. Moreover, none of the
aforementioned studies analyze the potential impact of delirium on the reliability and validity of observational pain instruments.

Finally, the symptom of pain is a complex, personal experience, making it difficult to adequately describe and treat without the patient’s self-report. Pain management providers frequently use describing words such as “burning” or “throbbing” to guide their management practices. Without the benefit of that type of subjective data, patients are at risk for receiving substandard pain control with a less effective pharmacologic agent (i.e. narcotics for the treatment of neuropathic pain). Future studies investigating strategies that promote self-report of pain in sedated and/or mechanically ventilated critically ill patient would contribute much to the advancement of pain assessment in this patient population.

**Conclusion**

The rates of uncontrolled pain in the critically ill remain unacceptably high, with an overwhelming majority of patients reporting moderate to severe levels of pain during their stay in the intensive care unit. 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 a validated pain assessment instrument has been shown to decrease length of stay, decrease the duration of mechanical ventilation, and increase patient, family, and provider satisfaction. Of the available observational pain scales, the CPOT has shown superior reliability and validity when utilized in this patient population. More research is indicated in the assessment of pain in individuals incapable of spontaneous neuromuscular movement and in those with chronic pain.
References


