State Behavioral Scale: A sedation assessment instrument for infants and young children supported on mechanical ventilation*

Martha A. Q. Curley, RN, PhD, FAAN; Sion Kim Harris, PhD; Karen A. Fraser, RN; Rita A. Johnson, RN, BSN; John H. Arnold, MD

Objective: **To develop and test the reliability and validity of the State Behavioral Scale for use in describing sedation/agitation levels in young intubated patients supported on mechanical ventilation.**

Design: **In this prospective, psychometric evaluation, pairs of trained pediatric critical care nurse evaluators simultaneously and independently assessed a convenience sample of pediatric intensive care unit patients along eight state/behavioral dimensions and a numeric rating scale (NRS) of 0 (extremely sedated) to 10 (extremely agitated). The eight dimensions were derived from the sedation/agitation literature and expert opinion and included respiratory drive, response to ventilation, coughing, best response to stimulation, attentiveness to careprovider, tolerance to care, consolability, and movement after consoled, each with 3–5 levels.**

Setting: **An 18-bed pediatric medical–surgical intensive care unit and 26-bed pediatric cardiovascular intensive care unit in a university-affiliated academic children's hospital.**

Patients: **A total of 91 intubated mechanically ventilated patients 6 wks to 6 yrs of age provided a median of two observations (interquartile range, 1–3) for a total of 198 sets of observations. Excluded were postoperative patients or those receiving neuromuscular blockade.**

Interventions: **Patients were observed for 1 min, and then incremental levels of stimulation were applied until patient response. After 2 mins of consoling, the state behavioral assessment and NRS were completed.**

Measurements: **Weighted kappa and intraclass coefficients were generated to assess interrater reliability of the eight dimension and NRS ratings. Distinct state behavior profiles were empirically identified from the dimension ratings using hierarchical cluster analysis using a squared Euclidean distance measure and between-groups linkage. Construct validity of these profiles was assessed by comparing group mean NRS scores using one-way analysis of variance.**

Main Results: **Weighted kappa scores for all 198 dimension ratings ranged from .44 to .76, indicating moderate to good interrater reliability. The intraclass coefficient of .79 was high for NRS ratings. Cluster analysis revealed five distinct state profiles, with mean NRS ratings of 1.1, 2.5, 4.0, 5.3, and 7.6, all of which differed significantly from each other ($F = 75.8$, $p < .001$), supporting the profiles' construct validity.**

Conclusions: **Based on empirically derived state behavior profiles, we have constructed the State Behavioral Scale to allow systematic description of the sedation–agitation continuum in young pediatric patients supported on mechanical ventilation. Further studies including prospective validation and describing the effect of State Behavioral Scale implementation on clinical outcomes, including the quality of sedation and length of mechanical ventilation, are warranted. (Pediatr Crit Care Med 2006; 7:107–114)**

KEY WORDS: sedation; agitation; pediatric intensive care; child; nursing assessment

Ensuring the comfort of critically ill infants and children is integral to the practice of pediatric critical care. Humane pediatric intensive care often includes the administration of sedatives after pain, physiologic imbalance, and environmental stressors have been addressed (1). More than 90% of infants and children supported on mechanical ventilation receive some form of sedative therapy (2). Sedation in this patient population is required for anxiolysis, amnesia, facilitation of care, patient safety in avoidance of adverse events, and for decreasing oxygen consumption. For most young patients supported on mechanical ventilation, the goal of sedation is to attain a calm but responsive state (2, 3) that protects the young patient from self-harm. Inadequate sedation is associated with potentially dangerous complications such as unplanned endotracheal extubation (4). High-dose, long-term, and continuous intravenous sedation has been associated with prolonged weaning from mechanical ventilation and withdrawal syndrome (5–7). Therefore, insufficient or excessive sedation is likely to add to the personal and financial burden of intensive care.

Variability complicates the use of sedation in the pediatric intensive care unit (ICU) setting. First, a patient's sedative needs vary depending on the nature and course of the illness, drug interaction with concomitant therapies, and response to therapy (8, 9). Some patients require deep sedation to tolerate synchronous modes of mechanical ventilation, whereas others seem to be comfortable with light sedation, even when supported on unconventional modes of ventilatory support (10). Next, from a systems per-

*See also p. 183.

From the Critical Care and Cardiovascular Program (MAQC), Clinical Research Program (SKH), Clinical Research Center (KAF), and Medical–Surgical Intensive Care Unit (RAJ, JHA), Children's Hospital Boston, Boston, MA.

Supported, in part, by NIH/NICHD 5R21HD045020-02 and GCRC MO1-RR02172 from the Gustavus and Louise Pfeiffer Research Foundation.

The authors have no financial interests to disclose.

DOI: 10.1097/01.PCC.0000200955.40962.38

spective, multidisciplinary staff with varying levels of expertise change several times over the working day. This means that patients are exposed to multiple subjective assessments of their sedation requirements, by several staff members, which may result in patients receiving varying dosages of sedation, depending on who performed the assessment (11, 12).

Valid and reliable tools that standardize the description of a pediatric patient's behavioral state while supported on mechanical ventilation would enhance systematic assessment and documentation of a patient's response to sedation, allow patient-specific alterations in the therapeutic regimen, and help avoid insufficient or excessive sedative use (13, 14). Such an assessment tool would enhance interdisciplinary agreement on the desired level of sedation, provide a foundation for the development of guidelines that would decrease unnecessary variation in the care (15–17), and permit objective study of the pharmacodynamics of sedative agents in the pediatric population (18). From a research perspective, Kollef et al. (5) suggest that sedation practices should be standardized in any investigation employing the duration of mechanical ventilation as an outcome variable.

Although desirable, tools assessing the sedation–agitation continuum in the pediatric patient have not been adequately tested or have conceptual flaws. Specifically, the psychometrics of the Ramsay scale (19), an often-cited sedation scale used in the adult population, has never been evaluated in an ICU setting. In addition, Ramsay's six levels of sedation are neither mutually exclusive nor clearly defined (20). The COMFORT scale (14, 21), the most commonly used tool in the pediatric population (12), was designed to assess distress in ventilated children, but distress was operationalized to include the constructs of both pain and agitation. Noting that the eight dimensions of the COMFORT scale are often included in other pain instruments, van Dijk et al. (22) supported the use of the COMFORT scale to assess postoperative pain in infants. From a clinical perspective, separate valid and reliable pain and agitation assessment tools would allow more targeted therapeutic management (23).

The purpose of this study was to empirically construct and demonstrate preliminary construct validity and interrater reliability of a pediatric sedation assessment scale, the State Behavioral Scale (SBS), for use in young critically ill pediatric patients supported on mechanical ventilation.

MATERIALS AND METHODS

Data Collection Instrument. A state behavioral assessment tool was derived from our previous work describing pediatric ICU nurses' descriptions of agitation (24), the literature on sedation tools used in the adult population (25–27), and expert opinion from a pediatric anesthesiologist and pediatric critical care clinical nurse specialist. Two adult ICU sedation scales, the Sedation–Agitation Scale (25) and, its derivative, the Motor Activity Assessment Scale (27, 28), served as templates. As presented in Figure 1, we retained descriptors that could be evaluated in a cognitively immature patient population and added several descriptors important to the care of young patents supported on mechanical ventilation. Our tool included ratings along the following eight dimensions: respiratory drive, response to ventilation, coughing, best response to stimulation, attentiveness to careprovider, tolerance to care, consolability, and movement after consoled. Each dimension contained three to six levels that incrementally described the sedation–agitation continuum.

Operational definitions included the following: sedation, a calm tranquil state that allays anxiety and excitement; agitation, excitement accompanied by increased motor activity; attention, the ability to open eyes and notice surroundings; response, the ability to open eyes, or raise eyebrows, or turn head toward stimulus, or move limbs; distress, sudden increase in heart rate or blood pressure, and/or a decrease in Spo_2, or increase in movement. The 0–10 numeric rating scale (NRS), with 0 equal to "extremely sedated" and 10 equal to "extremely agitated," served as a reference standard. We did not include changes in heart rate and blood pressure as distinct dimensions in the state behavioral assessment because of consistently low sensitivity and specificity in predicting agitation (21, 24, 29). We also did not include the patient's ability to communicate, follow commands, or attempts to sit or climb out of bed because these items are not consistently developmentally appropriate across the 6-wk-old to 6-yr-old age group.

Patient Sample. Between 2000 and 2004, we enrolled a convenience sample of patients, 6 wks to 6 yrs of age, who were intubated and mechanically ventilated in either the medical–surgical ICU or cardiovascular ICU in a university-affiliated academic children's hospital in the Northeast. Enrollment was stratified in three age groups (6 wks to 1 yr, 1 to 3 yrs, and 3 to 6 yrs) to ensure an almost equal distribution of age within our sample. We excluded patients receiving neuromuscular blockade, postoperative patients, patients assessed to be in pain by their bedside nurse, patients who were considered physiologically unstable (those experiencing any increase in ventilatory or vasopressors support in the previous 2 hrs), and patients at risk for opioid withdrawal. Exclusion criteria were selected to eliminate patients incapable of providing behavioral clues and patients more likely to provide pain-related clues. Data collection did not alter in any way the current practice or the administration of sedatives. The study was approved by the institutional review board, and need for informed consent was waived because data were de-identified, considered to be low-risk, and were collected during routine patient care.

Data Collection. A pair of trained pediatric critical care nurse evaluators simultaneously and independently conducted state behavioral assessments of each intubated, mechanically ventilated pediatric patient in the sample. As part of their training before study implementation, a total of five critical care nurses were instructed on the use of the state behavioral assessment tool and NRS over the study period. After instruction, each nurse conducted five concurrent assessments per protocol with the principal investigator. The group members then discussed their assessments and reached consensus on the best score within each dimension of the state behavioral assessment and NRS. Data collection commenced when the nurse raters agreed that they had reached a common understanding of the eight state behavioral dimensions and the NRS and when there was 95% agreement between the critical care nurse rater and principal investigator, with disagreements not crossing more than one level in any dimension. Training data were not included in the final analyses.

Patients were enrolled as soon as possible after meeting criteria and were assessed daily for a maximum of 6 days. Before data collection, patients were presumed to be supported on appropriate ventilatory settings. Data collection was conducted at a time when the bedside nurse was completing planned care when two evaluators were available. If endotracheal extubation was planned, the patient was assessed just before the procedure.

First, patients were observed undisturbed for 1 min. Second, the patient's nurse provided progressive stimuli, as necessary, to elicit a patient's response. Specifically, the nurse first spoke the patient's name using a calm voice and then, if there was no response, spoke the patient's name and gently touched the patient's body. If there was still no response, the patient's response to a planned noxious stimulus was assessed, such as endotracheal suctioning or <5 secs of nail-bed pressure. Finally, the patient was repositioned, then consoled by the nurse, parent, or by both. After 2 mins of consoling, evaluators completed the state behavioral assessment and gave an NRS rating. In evaluating the coughing dimension, the evaluator was allowed to query the bedside nurse about this dimension if a suctioning procedure was not observed. In addition to the behavioral ratings, demo-

Dimensions	Levels
Respiratory Drive	1. No spontaneous respiratory effort 2. Spontaneous but ineffective exhaled tidal volume (Patient specific: <4cc/kg) 3. Spontaneous and effective exhaled tidal volume (Patient specific: > 4cc/kg)
Response to Ventilation	1. No spontaneous breathing 2. Easy spontaneous breathing (fully synchronized with mechanical ventilation) 3. Having difficulty synchronizing with ventilator 4. Unsynchronized with mechanical ventilation, compromising oxygenation/ventilation
Coughing	1. No cough with suction 2. Coughs only when suctioned 3. Coughs when repositioned 4. Occasional spontaneous cough 5. Frequent spontaneous coughing that does not resolve with suctioning 6. Bronchospastic or choking
Best Response to Stimulation	1. No response to noxious stimuli 2. Responds to noxious stimulus 3. Responds to touch 4. Responds to voice 5. No external stimulus is required to elicit response
Attentiveness to Care Provider	1. Unable to pay attention to care provider 2. Able to pay attention to care provider but drifts off after stimulation 3. Spontaneously pays attention to care provider (Infant – fixes and follows) 4. Vigilant of care provider/Eyes follow care provider – watchful 5. Hyper-vigilant to care provider/Panicky when care providers approach patient
Tolerance to Care	1. Does not distress with any procedure including noxious 2. Will distress with noxious procedures 3. Distresses with procedures (i.e., repositioning) 4. Distressed (e.g., picking at tubes, pulling at restraints, etc.) 5. Patient intermittently unsafe (e.g., biting ETT) 6. Patient unsafe (e.g., attempting to pull at ETT/catheters, cannot be left alone)
Consolability	1. Self-regulates/modulates own behavior 2. Able to calm with comforting touch or voice when stimulus removed; distractible 3. Does not consistently calm despite a 5-minute attempt to console 4. Unable to console
Movement after Consoled	1. Does not move 2. Occasional movement of extremities or shifting of position in bed 3. Increased movement (restless, squirming) 4. Excessive movement (thrashing side to side, kicking legs, arched, rigid) 5. Combative

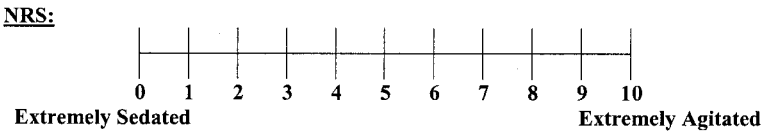

Figure 1. Behavioral assessment tool and numeric rating scale (*NRS*). *ETT*, endotracheal tube.

graphic data, including age, sex, admission functional health (30), admission PRISM III (31), and the use of mechanical ventilation and sedative use at the time of assessment, were recorded.

Statistical Analysis. We generated descriptive statistics (mean values and standard deviations, median values and interquartile ranges, response proportions) on sample patient demographic characteristics including age, sex, race/ethnic group, primary and secondary diagnoses, cerebral and overall performance, and risk of mortality. We assigned numeric codes to each level of each state behavioral dimension for analyses, with 1 for the first level, and a high of 3 to 6, depending on the number of levels within a dimension. After reviewing frequency distributions, the respiratory drive and response to ventilation dimensions were collapsed into one dimension for subsequent analyses. The dimensions were collapsed because few patients exhibited, under respiratory drive, spontaneous but ineffective exhaled tidal volume or were, under response to ventilation, unsynchronized with mechanical ventilation, compromising oxygenation and ventilation.

Interrater agreement for the seven dimension ratings, consisting of ordinal data, was assessed using weighted kappa coefficients generated with PROC FREQ in SAS (SAS Institute, Cary, NC) (32). We examined interrater agreement on the NRS rating, a continuous variable, with intraclass correlation coefficients generated in SPSS 12.0 (SPSS, Chicago, IL) using a one-way random-effects model. We analyzed interrater reliability using all available paired ratings, which included multiple rating occurrences for some patients. Because there may have been intracluster correlation among multiple ratings of the same patient, we also conducted these analyses using only the first rating of each patient and using a randomly selected single rating for each patient. To explore whether interrater agreement varied by patient age group, all interrater reliability analyses were stratified by the three patient age groups.

To identify distinct state behavior profiles based on dimension ratings, we conducted hierarchical cluster analysis using a squared Euclidean distance measure to assess similarity/dissimilarity across cases and a between-groups linkage method for combining clusters. We ran the cluster analysis using data from rater 1 only, rater 2 only, and data averaged across the two raters. In addition, we conducted these analyses using all rating occurrences and also using only data from a randomly selected single rating for each patient. Random selection of a single rating per patient was deemed preferable to using the first rating for each patient in these analyses because the first ratings tended to reflect more-sedated states compared with subsequent ratings. Patients were generally more sedated immediately after intubation.

Because the seven dimensions had differing response scale ranges, we ran the cluster analyses using standardized scores, in which item scores were converted into z-scores. We chose the number of clusters by examining the agglomeration schedule and identifying the "elbow" in the curve of the distance measure across the cluster-joining steps (the point at which the distance coefficient made a sudden jump in size). Although this method of identifying the number of clusters indicated that there were likely to be five clusters or state-behavior profiles, we evaluated three-, four-, and five-cluster solutions in subsequent

Table 1. Patient demographics

Characteristic	(n = 91)
Age, mos (median [interquartile range])	18 (4.4–34.5)
Age group, n (%)	
6 wks to 1 yr	38 (42)
1–3 yrs	31 (34)
3–6 yrs	22 (24)
Female sex, n (%)	34 (37)
Race or ethnic group, n (%)	
White	63 (69)
Black	4 (4)
Hispanic	11 (12)
Asian/Pacific Islander	3 (3)
More than one group	10 (11)
Pediatric Cerebral Performance Category,[a] median (first quartile, third quartile)	1 (1, 2)
Pediatric Overall Performance Category,[b] median (first quartile, third quartile)	1 (1, 3)
Pediatric Risk of Mortality III scores,[c] median (interquartile range)	6 (0–12)
Mortality	12 (13)
Primary diagnosis, n (%)	
Pulmonary	46 (51)
Cardiovascular	12 (13)
Neurologic	9 (10)
Infectious disease/sepsis	6 (7)
Congenital anomaly	6 (7)
Gastrointestinal	5 (6)
Oncology/bone marrow transplant	5 (6)
Metabolic/trauma	2 (2)
Secondary diagnoses, n (%)	45 (49)

[a]Pediatric Cerebral Performance Category score ranges from 1 (normal cognitive development) to 6 (brain death) (30); [b] Pediatric Overall Performance Category score ranges from 1 (good overall performance) to 6 (brain death) (30); [c] scores from the Pediatric Risk of Mortality III can range from 0 to 74, with higher scores indicating higher probability of death (31). Because of rounding, percentages may not total 100.

analyses to further determine the appropriateness of the five-cluster solution. After examining the median scores for each of the seven SBS dimensions across the clusters, we determined that the five-cluster solution was appropriate, with the five groups showing distinct profiles of scores across the seven state dimensions. To assess discriminant validity of the profiles, we used a one-way analysis of variance with a *post hoc* Tukey honest-significance-difference test, to compare mean scores on the NRS variable across the five cluster groups. A two-sided *p*-value of $<.05$ indicated statistical significance. All cluster and mean score comparison analyses were performed with SPSS 12.0 software.

Specific information needed (frequency distributions and correlations between raters) to perform power calculations for kappa coefficients was not known, leading to the need for preliminary data to be collected through this study. However, cross-sectional analyses on 90 subjects provides 82% power to detect interrater correlations of 0.30 and 98% power to detect correlations of 0.40. Thus, we anticipated that our sample size of 90 subjects would be sufficient to detect moderate to high correlations and kappa coefficients and would provide enough pilot data for us to develop appropriate follow-up studies.

RESULTS

A total of 91 patients were enrolled. Patient demographic data are presented in Table 1. Most patients were cognitively and functionally normal for age (30). The majority of the 91 patients (n = 57, 63%) were intubated and mechanically ventilated for pulmonary parenchymal disease, with the rest having airways disease. Although conventional methods were used to ventilate most patients, 9% (n = 8) were supported on high-frequency oscillatory ventilation and 3% were supported on extracorporal membrane oxygenation. Mortality rate was 13%. Almost all of the patients (95%) were receiving sedation. The most prevalent sedation plan (73%) included a combination of opioids and benzodiazepines. Pairs of nurse evaluators assessed the 91 patients during 198 rating occurrences (396 total observations). Most patients provided one (46%) or two (23%) sets of paired observations.

In analysis of interrater agreement across all 198 available paired ratings, weighted kappa coefficients ranged from .44 (consolability) to .76 (respiratory drive/response to ventilation) across the seven dimension ratings, indicating moderate to good interrater agreement (Table 2) (33). This conclusion was additionally supported by weighted kappas generated from analysis of data that included only the first rating for each patient (range, .55–.77) and of a randomly selected single rating for each patient (range, .54–.74; data not shown). The actual concordance rates for the seven SBS dimensions (percentage of ratings for which the two raters gave the exact same rating), using all 198 available ratings, ranged from 66% (tolerance to care) to 85% (respiratory drive/response to ventilation) (data not shown).

Similarly, interrater agreement for the NRS rating was good, with an intraclass coefficient of .79 across all 198 paired ratings, .83 using only data from the first ratings, and .79 using randomly selected single ratings.

The level of interrater agreement did not differ by patient age group, except for the coughing dimension. In analyses of all 198 paired ratings, the weighted kappa for this dimension was significantly lower in the youngest age group of 6 wks to 1 yr (.55; 95% confidence interval, .41–.69), compared with the weighted kappa in the oldest age group of 3–6 yrs (.81; 95% confidence interval, .68–.93). The middle age group, ages 1–3 yrs, had a weighted kappa of .73 (95% confidence interval, .68–.93).

In Table 2, we show the state behavioral profiles of the five groups identified from cluster analysis. We present the results of analysis of all 198 ratings, which did not differ from those of 91 randomly selected single-patient ratings. Also, because interrater agreement was generally good for the paired dimension ratings, we are presenting the results of cluster analysis using averaged data for the rater pairs. As shown by the median scores on each of the seven dimensions, and mean NRS ratings, these five groups had distinct profiles of scores. Specifically, these cluster groups had significantly different mean scores on the NRS (one-way analysis of variance $F = 75.8$, $df = 4$, $p < .001$), with each group differing from all other groups in *post hoc* pair-wise comparisons (all $p < .002$), supporting the discriminant validity of these five profiles.

Using the median scores on the seven dimensions and mean NRS scores for each group (Table 2), the five profiles were then aligned to a bipolar numeric scale ranging from −3 to +1. We selected this metric because our group had experience using a modified Motor Activity Assessment Scale (27) with a bipolar scale that clinicians described as logical and readily recalled; specifically, negative numbers equated to less-active states and more positive numbers equated to more-active states. This

Table 2. Median score (interquartile range [IQR]) on each State Behavior Scale (SBS) dimension and mean (95% confidence interval [CI]) numeric rating scale score (NRS), within each of five rating groups generated from cluster analysis using all ratings (n = 198); interrater reliability coefficients (weighted kappa and intraclass correlation [ICC]) for each SBS dimension and NRS using all ratings, and using only the first rating per patient (n = 91)

SBS Dimension	−3 (n = 33) Median (IQR)	−2 (n = 72) Median (IQR)	−1 (n = 26) Median (IQR)	0 (n = 59) Median (IQR)	+1/+2 (n = 8) Median (IQR)	Weighted Kappa ALL (95% CI)	Weighted Kappa FIRST (95% CI)
Respiratory drive/response to ventilation	No spontaneous respiratory effort 1.0 (1.0–1.5)	Spontaneous and effectively supported breathing 3.0 (3.0–3.0)	Spontaneous but ineffective/ effectively supported breathing 2.8 (1.9–3.0)	Spontaneous and effective breathing 3.0 (3.0 3.0)	Spontaneous and effective/having difficulty synchronizing with ventilator 3.0 (3.0–3.9)	.76 (.67–.84)	.77 (.65–.88)
Coughing	No cough/coughs only when suctioned 2.0 (1.0–2.0)	Coughs only when suctioned or when repositioned 3.0 (2.0–3.9)	Coughs only when suctioned or when repositioned 2.0 (2.0–3.0)	Coughs when repositioned/ occasional spontaneous cough 3.5 (3.0–4.0)	Occasional spontaneous cough 4.0 (4.0–4.0)	.68 (.59–.77)	.76 (.64–.87)
Best response to stimulation	No response to noxious stimuli 1.0 (1.0–2.0)	Responds to noxious stimuli/touch 2.0 (1.6–3.0)	Responds to touch/voice 3.5 (3.0–4.0)	Responds to voice/no external stimulus required to elicit response 4.5 (4.0–5.0)	Responds to voice/no external stimulus required to elicit response 4.8 (3.6–5.0)	.71 (.64–.78)	.65 (.54–.76)
Attentiveness to care provider	Unable to pay attention to care provider 1.0 (1.0–1.0)	Unable to pay attention to care provider 1.0 (1.0–1.0)	Able to pay attention to care provider but drifts off after stimulation 1.5 (1.0–2.0)	Spontaneously pays attention to care provider (infant fixes and follows) 3.0 (2.5–3.0)	Drifts off /spontaneously pays attention 2.5 (1.8–2.5)	.69 (.61–.76)	.67 (.56–.78)
Tolerance to care	Does not distress with any procedure (including noxious) 1.0 (1.0–1.0)	Will distress with noxious procedure 1.5 (1.0–2.0)	Distresses with procedures 2.5 (2.0–3.0)	Distresses with procedures 2.5 (2.0–3.0)	Intermittently unsafe 4.5 (4.0–5.0)	.63 (.55–.71)	.60 (.48–.73)
Consolability	Self-regulates/ modulates own behavior 1.0 (1.0–1.0)	Self-regulates/ modulates own behavior 1.0 (1.0–1.0)	Able to calm with comforting touch or voice when stimulus removed 1.5 (1.5–2.0)	Able to calm with comforting touch or voice when stimulus removed 2.0 (1.5–2.0)	Does not consistently calm despite 5-min attempt to console 2.8 (2.5–3.4)	.44 (.32–.55)	.62 (.46–.77)
Movement after consoled	Does not move 1.0 (1.0–1.0)	Does not move/ occasional movement of limbs or shifting of position 1.0 (1.0–1.5)	Occasional movement of limbs or shifting of position 2.0 (2.0–2.0)	Occasional movement/ increased movement (restless, squirming) 2.0 (2.0–2.5)	Increased movement (restless, squirming) 3.0 (2.6–3.0)	.61 (.52–.70)	.55 (.43–.68)
Mean NRS (95% CI)	1.1 (0.7–1.6)	2.5 (2.1–2.9)	4.0 (3.4–4.5)	5.3 (4.9–5.6)	7.6 (6.9–8.4)	ICC = .79 (.73–.84)	ICC = .83 (.76–.89)

bipolar NRS was also used in the original Sedation–Agitation Scale for critically ill adults (26). The first cluster group with a mean (95% confidence interval) NRS score of 1.1 (0.7–1.6) was linked to an SBS dimension of −3, whereas the fourth cluster group with a mean NRS score of 5.3 (4.9–5.6) was equated with a SBS dimension of 0. The remaining −2, −1, and +1 values were then aligned around these fixed points.

Table 3 presents the empirically generated pediatric SBS with standardized definitions given for each point on the scale, similar to those used in the adult Motor Activity Assessment Scale (27). We also added a +2 agitated level because, although rarely observed, and thus not empirically captured in this study, experienced clinicians have cared for patients exhibiting these agitated behaviors in the pediatric ICU.

DISCUSSION

We empirically constructed a standardized SBS to describe the sedation–agitation continuum in an extremely vulnerable patient sample of young pediatric patients supported on mechanical ventilation. The SBS was derived from ratings of seven content dimensions (originally eight, but two dimen-

Table 3. State Behavioral Scale score as patient's response to voice, then gentle touch, then noxious stimuli (planned endotracheal suctioning or <5 secs of nail-bed pressure)

Score	Description	Definition
−3	Unresponsive	No spontaneous respiratory effort No cough or coughs only with suctioning No response to noxious stimuli Unable to pay attention to care provider Does not distress with any procedure (including noxious) Does not move
−2	Responsive to noxious stimuli	Spontaneous yet supported breathing Coughs with suctioning/repositioning Responds to noxious stimuli Unable to pay attention to care provider Will distress with a noxious procedure Does not move/occasional movement of limbs or shifting of position
−1	Responsive to gentle touch or voice	Spontaneous but ineffective nonsupported breaths Coughs with suctioning/repositioning Responds to touch/voice Able to pay attention but drifts off after stimulation Distresses with procedures Able to calm with comforting touch or voice when stimulus removed Occasional movement of limbs or shifting of position
0	Awake and able to calm	Spontaneous and effective breathing Coughs when repositioned/occasional spontaneous cough Responds to voice/no external stimulus is required to elicit response Spontaneously pays attention to care provider Distresses with procedures Able to calm with comforting touch or voice when stimulus removed Occasional movement of limbs or shifting of position/increased movement (restless, squirming)
+1	Restless and difficult to calm	Spontaneous effective breathing/having difficulty breathing with ventilator Occasional spontaneous cough Responds to voice/no external stimulus is required to elicit response Drifts off/spontaneously pays attention to care provider Intermittently unsafe Does not consistently calm, despite 5-min attempt/unable to console Increased movement (restless, squirming)
+2	Agitated	May have difficulty breathing with ventilator Coughing spontaneously No external stimulus required to elicit response Spontaneously pays attention to care provider Unsafe (biting endotracheal tube, pulling at catheters, cannot be left alone) Unable to console Increased movement (restless, squirming, or thrashing side-to-side, kicking legs)

sions were combined) derived from the literature and expert opinion, and we describe adequate content and construct validity and interrater reliability of these dimensions. An ideal sedation scale for pediatric intensive care should be valid and reliable, be developmentally appropriate, integrate the multidimensional goals of sedation, be easy to complete and interpret, contain precise discriminating criteria at each level, and be useful in directing sedative therapy (34–36). We believe the SBS meets these criteria.

The SBS was specifically designed for and tested in young, intubated, mechanically ventilated patients—a population of patients who often experience extremes in level of sedation over their normal trajectory of illness. The tool reflects the presence and severity of the clinical conditions for which sedation is administered in this population (20, 36). We designed the SBS to require an evaluation of patient response to a progressive stimulus. This approach is familiar to intensive care clinicians as it is similar to the progressive stimulus performed during a neurologic examination. We believe that appropriate levels of sedation should help the young patient through necessary care procedures that require a stimulated state while avoiding oversedation when not stimulated.

The dimensions include descriptors that have been previously described to be associated with agitation (24). More negative scores reflect a more-sedated state. More positive scores reflect a more agitated state. The single-digit bipolar numeric avoids the complexity of summing multiple dimension scores and is logical in that the use of negative numbers for sedation, positive numbers for agitation, and zero score for neither sedation nor agitation may enhance clinician recall of the measure. Each level contains multiple descriptors, increasing the likelihood that a patient's behavior can be mapped to a single level. The SBS was developed to augment a clinician's clinical judgment. Differentiating behavioral distress from physiologic distress requires the clinician to interpret patient behavior within the context of an evolving clinical state.

Although the ease with which the nurse evaluators rated each dimension was not systemically described, the overall experience of the raters was that the dimensions were clear and that completion of the behavioral assessment tool and NRS was easily accomplished in <2 mins after the stimulation protocol was implemented. However, some dimensions may be more challenging to rate

than others; for example, a patient may console differently to a nurse's voice/touch and to a known caregiver's voice/touch. The relatively low weighted kappa reported for consolability may also be in part due to a sensitivity of the kappa to the response distribution, which in the case of this item, was highly skewed. Over 95% of ratings by either rater consisted of "self-regulates" or "able to calm; distractible," with only 4% giving a rating of "does not consistently calm" or "unable to console." The actual concordance rate between raters for this item was moderately high, at 69% among all 198 ratings (79% among first ratings, 74% among randomly selected single ratings), indicating interrater agreement comparable with the other dimensions.

Except for the coughing dimension, the level of interrater agreement did not differ by patient age group. The weighted kappa for coughing was significantly lower in the youngest age group (6 wks to 1 yr) compared with the weighted kappa in the oldest age group (3–6 yrs). It should be noted that coughing was the only dimension that was not always directly observed. We are unable to determine in this study whether coughing is a more salient dimension for certain age groups compared with others, thus affecting interrater agreement across the age groups. However, interrater agreement of the coughing dimension may become more consistent across age groups when bedside nurses directly observe the phenomenon. Future studies evaluating the new SBS tool should assess the ease/difficulty in which the rating is made.

It is not surprising that an extreme agitated state was not well observed in this patient sample. Although these behaviors do occur, we hypothesize that nurses immediately intervene to manage evolving unsafe and inconsolable behaviors in young, intubated, mechanically ventilated patients. Our findings are similar to those of Sessler et al. (37), who validated the Richmond Agitation-Sedation Scale in a wide distribution of critically ill adult patients. The Richmond Agitation-Sedation Scale ranges from −5 (unresponsive) to +4 (combative). Only 10% of their observations were in the +1-restless to +3–very agitated range, and none was noted to be +4 combative.

We recommended that sedation assessment be completed at the start of normal cares at a frequency that aligns with the patient's clinical state. The Joint Commission for the Accreditation of Hospitals suggests that pain scoring be considered the fifth vital sign (38). We suggest that sedation scoring should be completed with the patient's pain assessment every 4 hrs and also before and after an intervention that affects the patient's level of sedation.

The use of a convenience sample was a limitation of this study and may reduce the generalizability of our findings. Data collection required the presence of two trained observers, which precluded consecutive sampling. Data collection also did not alter current practice or the administration of sedatives; thus, few patients exhibited an agitated state. Specifically, seven patients accounted for a total of eight +1/+2 SBS ratings (one patient accounted for two ratings). Our data indicated that nurses were successful in keeping their patients in a more-sedated state while intubated. Toward the end of the data collection period, we attempted to identify more-awake patients by focusing on the enrollment of patients just before endotracheal extubation. Next, although we excluded patients assessed to be in pain, we cannot be completely certain that these developmentally nonverbal or verbal intubated patients were pain-free. Pain and sedation scoring require clinical judgment and an evaluation of the context of patient trajectory and history. Future studies should test the construct validity of this instrument compared with a valid and reliable pain tool. Future studies should also assess the sensitivity of the SBS to assess a change in a patient's state over time or after sedative administration to determine its usefulness in informing patient-specific alterations in the therapeutic regimen and to describe the effect of illness/injury on the patient's state behavior. Finally, because the SBS tool was developed using ratings of patients aged 6 wks to 6 yrs who were physiologically stable and not rated to be in pain, other studies are needed to assess the appropriateness of this tool for use in populations not included in this study. Additional studies are also needed, using much larger samples, to assess the validity of the SBS profiles within each of the three pediatric age groups included in this study. Although we had equal representation of the three age groups in our sample, we had too few observations in the agitated range to be able to conduct cluster analyses stratified by age group.

In summary, of primary concern to all clinicians caring for critically ill pediatric patients is to limit the negative impact of the illness on the developing child. Many factors contribute to the process of providing this humanistic element of care, one of which includes ensuring an adequate level of sedation. We believe that the pediatric ICU environment taxes the adaptive capacities of even our most resilient patients and that our preverbal patients are at a particular disadvantage because they are too cognitively immature to process the importance of tolerance to invasive procedures, instrumentation, and support. The SBS describes the sedation–agitation continuum in this vulnerable group.

ACKNOWLEDGEMENTS

We thank the pediatric critical care nurses and our patients and their families who supported this study; David Wypij, PhD, who provided comments early in the design of the study; and Michelle Labrecque, MSN, RN, and Patricia Jones, MSN, RN, who assisted with patient screening and data collection.

REFERENCES

1. Gordin PC: Assessing and managing agitation in a critically ill infant. *MCN Am J Matern Child Nurs* 1990; 15:26–32
2. Marx CM, Rosenberg DI, Ambuel B, et al: Pediatric intensive care sedation: Survey of fellowship training programs. *Pediatrics* 1993; 91:369–378
3. Durbin CG: Sedation in the critically ill patient. *New Horiz* 1994; 2:64–74
4. Chevron V, Menard JF, Richard JC, et al: Unplanned extubation: Risk factors of development and predictive criteria for reintubation. *Crit Care Med* 1998; 26:1049–1053
5. Kollef MH, Levy NT, Ahrens TS, et al: The use of continuous i.v. sedation is associated with prolongation of mechanical ventilation. *Chest* 1998; 114:541–548
6. Fonsmark L, Rasmussen YH, Carl P: Occurrence of withdrawal in critically ill sedated children. *Crit Care Med* 1999; 27:196–199
7. Randolph AG, Wypij D, Venkataraman ST, et al: Effect of mechanical ventilator weaning protocols on respiratory outcomes in infants and children: A randomized controlled trial. *JAMA* 2002; 288:2561–2568
8. Playfor SD, Thomas DA: Quality of sedation during mechanical ventilation. *Paediatr Anaesth* 2000; 10:195–199
9. Young C, Knudsen N, Hilton A, et al: Sedation in the intensive care unit. *Crit Care Med* 2000; 28:854–866
10. Simmons LE, Riker RR, Prato BS, et al: Assessing sedation during intensive care unit

mechanical ventilation with the Bispectral Index and the Sedation–Agitation Scale. *Crit Care Med* 1999; 27:1499–1504
11. Westcott C: The sedation of patients in intensive care units: A nursing review. *Intensive Crit Care Nurs* 1995; 11:26–31
12. Twite MD, Rashid A, Zuk J, et al: Sedation, analgesia, and neuromuscular blockade in the pediatric intensive care unit: Survey of fellowship training programs [see comment]. *Pediatr Crit Care Med* 2004; 5:521–532
13. Hansen-Flaschen JH, Brazinsky S, Basile C, et al: Use of sedating drugs and neuromuscular blocking agents in patients requiring mechanical ventilation for respiratory failure: A national survey. *JAMA* 1991; 266: 2870–2875
14. Marx CM, Smith PG, Lowrie LH, et al: Optimal sedation of mechanically ventilated pediatric critical care patients. *Crit Care Med* 1994; 22:163–170
15. Bair N, Bobek MB, Hoffman-Hogg L, et al: Introduction of sedative, analgesic, and neuromuscular blocking agent guidelines in a medical intensive care unit: Physician and nurse adherence. *Crit Care Med* 2000; 28: 707–713
16. Brook AD, Ahrens TS, Schaiff R, et al: Effect of a nursing-implemented sedation protocol on the duration of mechanical ventilation. *Crit Care Med* 1999; 27:2609–2615
17. Kress JP, Pohlman AS, O'Connor MF, et al: Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation [see comment]. *N Engl J Med* 2000; 342:1471–1477
18. Ostermann ME, Keenan SP, Seiferling RA, et al: Sedation in the intensive care unit: A systematic review. *JAMA* 2000; 283: 1451–1459
19. Ramsay MA, Savege TM, Simpson BR, et al: Controlled sedation with alphaxalone-alphadolone. *BMJ* 1974; 2:656–659
20. Hansen-Flaschen J, Cowen J, Polomano RC: Beyond the Ramsay scale: Need for a validated measure of sedating drug efficacy in the intensive care unit. *Crit Care Med* 1994; 22:732–733
21. Ambuel B, Hamlett KW, Marx CM, et al: Assessing distress in pediatric intensive care environments: The COMFORT scale. *J Pediatr Psychol* 1992; 17:95–109
22. van Dijk M, de Boer JB, Koot HM, et al: The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0- to 3-year-old infants. *Pain* 2000; 84:367–377
23. Wansbrough SR, White PF: Sedation scales: Measures of calmness or somnolence? *Anesth Analg* 1993; 76:219–221
24. Curley M, McDermott B, Berry P, et al: Nurses decision making regarding the use of analgesics and sedatives in the pediatric ICU. *Heart Lung* 1992; 21:296
25. Riker RR, Picard JT, Fraser GL: Prospective evaluation of the Sedation-Agitation Scale for adult critically ill patients. *Crit Care Med* 1999; 27:1325–1329
26. Riker RR, Fraser GL, Cox PM: Continuous infusion of haloperidol controls agitation in critically ill patients. *Crit Care Med* 1994; 22:433–440
27. Devlin JW, Boleski G, Mlynarek M, et al: Motor Activity Assessment Scale: A valid and reliable sedation scale for use with mechanically ventilated patients in an adult surgical intensive care unit. *Crit Care Med* 1999; 27: 1271–1275
28. Clemmer TP, Wallace JC, Spuhler VJ, et al: Origins of the Motor Activity Assessment Scale score: A multi-institutional process. *Crit Care Med* 2000; 28:3124
29. Carnevale FARNP, Razack SMD: An item analysis of the COMFORT scale in a pediatric intensive care unit. *Pediatr Crit Care Med* 2002; 3:177–180
30. Fiser DH: Assessing the outcome of pediatric intensive care. *J Pediatr* 1992; 121:68–74
31. Pollack MM, Patel KM, Ruttimann UE: The Pediatric Risk of Mortality III–Acute Physiology Score (PRISM III-APS): A method of assessing physiologic instability for pediatric intensive care unit patients. *J Pediatr* 1997; 131:575–581
32. Cohen J: Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968; 70:213–220
33. Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159–174
34. De Jonghe B, Cook D, Appere-de-Vecchi C, et al: Using and understanding sedation scoring systems: A systematic review. *Intensive Care Med* 2000; 26:275–285
35. Chulay M: Sedation assessment: Easier said then done! *Crit Care Nurs Clin North Am* 2004; 16:359–364
36. Lieberman J, Tremper KK: Sedation: If you do not know where you are going, any road will get you there. *Crit Care Med* 1999; 27: 1395–1396
37. Sessler CN, Gosnell MS, Grap MJ, et al: The Richmond Agitation-Sedation Scale: Validity and reliability in adult intensive care unit patients. *Am J Respir Crit Care Med* 2002; 166:1338–1344
38. Kerns RD, Wasse L, Ryan B: Pain as the 5th Vital Sign Toolkit. Vol. V. 2. Washington, DC, Veterans Health Administration, 2000