Validation of the Pediatric Sedation State Scale

Joseph P. Cravero, MD, a Nissa Askins, MPH, a Patcharee Sriswasdi, MD, b Daniel S. Tsze, MD, MPH, b David Zurakowski, PhD, c Sean Sinnott, MS a

OBJECTIVES: Development and validation of the Pediatric Sedation State Scale (PSSS) is intended to specifically meet the needs of pediatric procedural sedation providers to measure effectiveness and quality of care.

METHODS: The PSSS content was developed through Delphi methods utilizing leading pediatric sedation experts and published guidelines on procedural sedation in children. Video clips were created and presented to study participants, who graded the state of patients during procedures by using the PSSS to evaluate inter- and intrarater reliability by determining the intraclass correlation coefficient. We also compared the PSSS to the Observational Scale of Behavioral Distress–revised during 4 clinically relevant phases of a laceration repair procedure.

RESULTS: Six sedation states were defined for the PSSS. Each state was assigned a numerical value with higher numbers for increasing activity states. We included behaviors associated with adequate and inadequate sedation and adverse events associated with excessive sedation. Analysis of interrater and intrarater reliability revealed an intraclass correlation coefficient of 0.994 (95% confidence interval: 0.986–0.998) and 0.986 (95% confidence interval: 0.970–0.995), respectively. Criterion validity was confirmed with respect to the Observational Scale of Behavioral Distress–revised (Spearman r = 0.96). Construct validity was indicated by significant differences in PSSS scores (P < .001) between 4 phases of a procedure, each having a different degree of painful or distressing stimuli.

CONCLUSIONS: The PSSS is a 6-point scale that is a valid measure of the effectiveness and quality of procedural sedation in children within the limits of the testing method used in this study.

WHAT'S KNOWN ON THIS SUBJECT: Assessing the quality of sedation care is difficult. The literature includes scales that attempt to codify sedation level or pain behaviors but lacks a scale that describes the full range of possible patient conditions that can exist during a procedure.

WHAT THIS STUDY ADDS: We present the development and initial validation of the Pediatric Sedation State Scale, which is specifically intended to meet the needs of procedural sedation providers. The scale includes behaviors and adverse events that accompany procedural sedation.


DOI: 10.1542/peds.2016-2897

Accepted for publication Feb 2, 2017
Pediatric procedural sedation is provided by a wide range of professionals using many techniques and medications. The challenges facing procedural sedation providers vary widely with the type of procedure to be performed and the nature of the patient. Research attempting to assess the quality and safety of sedation practice has suffered from the lack of an objective scale that specifically allows the comparison of the efficiency, efficacy, and safety of sedation provided by various strategies and systems. The ideal procedural sedation scale would measure a variety of factors, such as control of movement, stress, and pain, while at the same time allowing for categorization of respiratory depression or other effects related to sedation. Importantly, this scale should be simple enough to be applied in a time frame that is appropriate for clinical procedural sedation. Scales that are currently used largely focus on the depth of sedation during procedures. Unfortunately, this measure does not actually reflect the quality of care with respect to other aspects of quality sedation. In addition, ICU pain scales that measure the degree of discomfort and stress of pediatric patients are often used to assess the management of pain during a procedure. These scales do not account for specific features that are relevant to providing the best possible patient state during brief painful/stressful interventions or diagnostic procedures.

Our research group previously conceived, and validated, the Dartmouth Operative Conditions Scale (DOCS; Supplemental Table 5), which was designed to facilitate video analysis of sedation techniques. The scale was relatively complex and gauged patient conditions in 4 dimensions at any given time point during a procedure using video analysis. With the Pediatric Sedation State Scale (PSSS), we sought to develop a scale to codify the appropriateness of the conditions present during sedated medical procedures for children in real time. Unlike the DOCS, the PSSS was intended to be simple and linear, allowing use during routine patient sedation for procedures. Furthermore, the scale is intended to be applicable to patients undergoing any sedation or analgesia technique (ie, simple distraction, sedation or anesthesia, or pure analgesia).

In this article, we present the methods that were used to develop the PSSS and validate its use as a tool in evaluating interventions used to decrease pain and anxiety.

METHODS

Model Development

In developing the PSSS, we used techniques commonly applied in the field of human factors engineering and similar to our previous experience developing the DOCS. This method begins with the theory that complex systems (like pediatric sedation delivery) are characterized by their input (patient before a procedure) and output (patient after a procedure) and the transfer function that relates the two. Modern control theory assumes an intermediate variable exists, in this case the state during the procedure. We defined the state of the patient as the overall condition of the patient including the experience of pain, anxiety, level of consciousness, and the presence of any adverse side effects from sedation (such as airway obstruction). By defining state, the operator will be better able to identify and control problem states and direct care toward a desirable state.

To ensure a comprehensive model of patient states for pediatric sedation, 3 tasks were completed:

1. The guidelines on pediatric sedation issued by the American Academy of Pediatrics and the American Society of Anesthesiology were used to help us define possible states and standard monitoring arrays (components of control). They were also used as a starting point to define the standard time course of procedures in which sedation was given;

2. A group of 9 procedural sedation experts who actively provide pediatric sedation and are involved in research relating to pediatric sedation were interviewed (Table 1). Input was sought by using Delphi methods. The range of possible patient states were identified as (1) unresponsive, apneic, (2) cyanotic, (3) hypotensive, (4) sleeping, (5) calm awake, (6) crying, (7) thrashing about, and (8) struggling vigorously. This range of patient states was recognized to be dynamic and changing over time. They also identified key milestones and goals that are common to all procedural sedations. Goals identified included control of anxiety, movement, adverse memories, and physiologic disturbance; and

3. The expert panel reviewed the data from the Pediatric Sedation Research Consortium, which included specific information on adverse events such as airway obstruction, apnea, inadequate sedation, prolonged recovery, and 25 other categories of events to orient them to the possible range of patient states that can occur during procedural sedation. A rating scheme was devised from the spectrum of patient states identified during development of the pediatric sedation model. Numerical values were assigned in such a way to reflect increasing numbers for increasing activity (decreasing sedation) states. The resulting scale was named the PSSS (Table 2).
Creation of Video Clips

After the design of the PSSS and institutional review board approval, 25 procedural sedations (from a variety of locations in each hospital) were videotaped from the time sedation was administered to the time patients were considered to be back to their baseline consciousness. The camera was placed in the most unobtrusive location possible that also allowed visualization of the patient and monitors. For this phase of the study, providers of sedation or anesthesia included radiology nurses, pediatricians, pediatric emergency medicine specialists, pediatric dentists, pediatric intensivists, and anesthesiologists. These videotapes represented over 190 hours of sedation.

Out of these videotape data, 40 video clips of 15-second duration were created to give short segments that included both a clear view of the patient and a simultaneous split-screen view of the patient monitor. Thirteen clips were then selected to ensure a range of patient states including (1) calm and alert, (2) crying/thrashing, (3) general anesthesia, (4) deep sedation, (5) minimal sedation, and (6) excessive sedation. These clips were used to evaluate interrater and intrarater reliability. We selected 50% of the video clip examples of calm, sedated patients and then continued to randomly select clips until all possible PSSS scores were included. A listing of the patient age and procedure for each clip is included in Supplemental Table 6.

Evaluating Interrater and Intrarater Reliability

The PSSS was tested on 20 sedation providers to assess interrater reliability and subsequently 6 months later (on a subgroup) to confirm intrarater reliability. These providers were chosen to reflect the variety of providers who might use the scale, including 4 anesthesiologists; 3 fellows in anesthesiology; 2 staff dentists; 2 nurse practitioners; 1 preoperative nurse; 1 radiology nurse; 4 recovery room nurses; 2 recovery room nurse assistants; and 1 participant who declined to specify her role. All raters were given a 5-minute Power Point presentation on the purpose and the use of the PSSS. Each rater was then presented with the 13 video clips and asked to rate them using a computerized PSSS rating screen. Scores were recorded electronically on the screen. To document intrarater validity, we retested 8 of the raters (who had been part of the original rater group) 6 months after their original rating by using the same 13 video clips and computer application (Fig 1). The retest group included 3 attending physicians in anesthesiology, 2 fellows in anesthesiology, 2 attending physicians in dentistry, and 1 recovery room nurse assistant.

Concurrent Criterion Validity

To demonstrate criterion validity, we determined the correlation between the PSSS and the Observational Scale of Behavioral Distress–revised (OSBD-r). The OSBD-r is an 8-factor, weighted observational scale used to measure distress associated with medical procedures validated in children 1 to 20 years of age (Supplemental Table 7). The total OSBD-r score is the sum of the OSBD-r scores for each factor, with a total score from 0 to 23.5 (0 = no distress, 23.5 = maximum distress).

For this phase of the study, OSBD-r and PSSS scores were obtained during a randomized clinical trial comparing 3 different volumes of intranasal midazolam in children undergoing laceration repair. After obtaining institutional review board approval (clinicaltrials.gov

<table>
<thead>
<tr>
<th>TABLE 1 Expert Panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel Member</td>
</tr>
<tr>
<td>Joseph P. Cravero, MD</td>
</tr>
<tr>
<td>Michael Mallory, MD</td>
</tr>
<tr>
<td>John Berkenbosch, MD</td>
</tr>
<tr>
<td>Marc Leder, MD</td>
</tr>
<tr>
<td>Gordon Gale, MD</td>
</tr>
<tr>
<td>Mark Roback, MD</td>
</tr>
<tr>
<td>Lynne Maxwell, MD</td>
</tr>
<tr>
<td>Baruch Krauss, MD</td>
</tr>
<tr>
<td>George T. Blike, MD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2 PSSS State</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Patient is moving (purposefully or nonpurposefully) in a manner that impedes the proceduralist and requires forceful immobilization. This includes crying or shouting during the procedure, but vocalization is not required. Score is based on movement.</td>
</tr>
<tr>
<td>4</td>
<td>Moving during the procedure (awake or sedated) that requires gentle immobilization for positioning. May verbalize some discomfort or stress, but there is no crying or shouting that expresses stress or objection.</td>
</tr>
<tr>
<td>3</td>
<td>Expression of pain or anxiety on face (may verbalize discomfort), but not moving or impeding completion of the procedure. May require help positioning (as with a lumbar puncture) but does not require restraint to stop movement during the procedure.</td>
</tr>
<tr>
<td>2</td>
<td>Quiet (asleep or awake), not moving during procedure, and no frown (or brow furrow) indicating pain or anxiety. No verbalization of any complaint.</td>
</tr>
<tr>
<td>1</td>
<td>Deeply asleep with normal vital signs, but requiring airway intervention and/or assistance (eg, central or obstructive apnea, etc).</td>
</tr>
<tr>
<td>0</td>
<td>Sedation associated with abnormal physiologic parameters that require acute intervention (ie, oxygen saturation &lt;90%, blood pressure is 30% lower than baseline, bradycardia receiving therapy).</td>
</tr>
</tbody>
</table>
registration number NCT01948908) and informed consent, 99 children between the ages of 1 and 7 with simple lacerations who required intranasal midazolam to facilitate their repair were enrolled. Procedures were videotaped for 1 minute before administration of the sedative (baseline) and from the time of administration of the sedative until 1 minute after completion of the laceration repair. Ninety-six patients had OSBD-r and PSSS scores available for analysis. Three blinded emergency medicine physicians independently scored each video in 15-second intervals during 4 clinically relevant phases: (1) baseline, (2) irrigation/cleaning, (3) laceration repair, and (4) postprocedure. Intranasal midazolam was administered after the baseline phase and before irrigation/cleaning. Each assessor was trained in the use of both the PSSS and the OSBD-r. During each 15-second interval, the child on the video was scored by using both the OSBD-r and the PSSS, allowing for comparison between the 2 measures. A total of 7057 comparisons were made.

**Construct Validity**

To evaluate construct validity, the same video clips (and their ratings) used to establish criterion validity were analyzed during the 4 clinically relevant phases: (1) baseline, (2) irrigation/cleaning, (3) laceration repair, and (4) postprocedure. Once again, 7057 individual measurements for 96 patients were considered: 1152 during preprocedure, 1492 during irrigation/cleaning, 3320 during laceration repair, and 1092 during postprocedure. The measurements occurred at 15-second intervals over all 4 of the clinical phases for up to 2 minutes for the irrigation/cleaning phase and up to 3 minutes for the laceration repair phase.

**Statistical Methods**

Interrater and intrarater (6 months later) reliability of the PSSS scale were measured by using the intraclass correlation coefficient (ICC) with 95% confidence interval (CI). Power analysis indicated that 20 rates and 13 video clips provided over 80% statistical power to estimate the ICC within a precision of 0.1. Criterion-related validity was assessed by using the Pearson correlation coefficient (r) for determining correlation between the ordinal PSSS scale (0–5) and the OSBD-R (0–23.5). Construct validation was evaluated by using the nonparametric Mann–Whitney U test to ascertain whether the PSSS scores varied according to different phases of surgical laceration repair, thereby establishing a relationship between changes in score across a range of procedural events. Box-and-whisker plots are used to summarize the median and interquartile range for PSSS scores. Two-tailed values of \( P < .05 \) were considered statistically significant. IBM SPSS Statistics was used for data analysis (SPSS version 23.0; IBM Corporation, Armonk, NY).

**RESULTS**

**Interrater Reliability**

Scores were compared between the 20 raters for each video clip. There was excellent agreement between
raters across videos as demonstrated in Table 3 (ICC = 0.994, 95% CI: 0.986–0.998, P < .001).

**Intrarater Reliability**

The scores for the video clips were compared for 8 raters between their initial rating and their rating 6 months later. The selection of raters who were retested was random. There was excellent agreement between the 2 viewings across the videos, so the rating did not differ according to the individual rater (Table 4) (ICC = 0.986, 95% CI: 0.970–0.995, P < .001).

**Criterion Validity**

Data from 96 separate patients were compared between the OSBD-r and the PSSS. These observations spanned the baseline, irrigation/cleaning, laceration repair, and postprocedure phases. Six patients did not have OSBD-r scores for all 4 phases, but all patients had scores for at least 1 phase. During each of the phases, there was a high correlation between the 2 scales (r = 0.841, P < .001), which is consistent with strong criterion validity (Fig 2).

**Construct Validity**

We compared PSSS scores during the 4 different clinically relevant phases. These phases included baseline, irrigation/cleaning, laceration repair, and the postprocedure. The PSSS scale varied based on the presence or absence of a distressing stimulus in each phase of the procedure as a whole, as would be expected if the scale was measuring changes in behavior related to a patient’s sedation state. Baseline scores increased during irrigation/cleaning and laceration repair phases of the procedure. They returned to close to baseline after all distressing stimuli were discontinued (Fig 3). The changes between the 4 phases were statistically significant (P < .001).

**DISCUSSION**

Clinical documentation of procedural sedation is often challenging because there is no scale that collects the primary, important outcomes from this practice in 1 simple tool. Similarly, results from pediatric procedural sedation trials are often difficult to compare because of the lack of a generally accepted tool for comparing the quality and safety of the sedation provided by a given drug or general strategy intervention. This PSSS was developed and validated with the goal of providing a single, simple scale that would specifically measure aspects of procedural sedation relating to the quality of sedation provided, including the control of pain, anxiety, movement, and adverse side effects.

We propose the use of this scale may help procedural sedation providers determine how often they achieve the ideal sedation state, which we define as a pain- and anxiety-free patient who is not moving and has normal vital signs during the procedure (PSSS Level 2).

We tested the PSSS in several dimensions. In terms of intrarater reliability, there was acceptable agreement in PSSS scores determined by different health care professionals (physicians and nurses) with a spectrum of professional training and levels of experience. Teaching the scale took no more than 10 minutes for each individual involved. In

**TABLE 3 Intrarater Reliability: PSSS Scores for Each Video Compared Between the 20 Raters Who Viewed and Scored the Video Clips**

<table>
<thead>
<tr>
<th>Video Clip No.</th>
<th>Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 (0–0)</td>
</tr>
<tr>
<td>2</td>
<td>2 (1–2)</td>
</tr>
<tr>
<td>3</td>
<td>1 (1–2)</td>
</tr>
<tr>
<td>4</td>
<td>2 (2–2)</td>
</tr>
<tr>
<td>5</td>
<td>2 (2–2)</td>
</tr>
<tr>
<td>6</td>
<td>2 (2–2)</td>
</tr>
<tr>
<td>7</td>
<td>2 (2–2)</td>
</tr>
<tr>
<td>8</td>
<td>2 (1–2)</td>
</tr>
<tr>
<td>9</td>
<td>3 (3–3)</td>
</tr>
<tr>
<td>10</td>
<td>4 (4–4)</td>
</tr>
<tr>
<td>11</td>
<td>5 (4–5)</td>
</tr>
<tr>
<td>12</td>
<td>5 (5–5)</td>
</tr>
<tr>
<td>13</td>
<td>5 (5–5)</td>
</tr>
</tbody>
</table>

Excellent agreement between the first and the second viewing for each of the 20 raters across the 13 videos. (ICC = 0.994, 95% CI: 0.985–0.998, P < .001). IQR, interquartile range.
evaluating intrarater reliability, the PSSS scores determined by a subset of the same health care professional at 2 separate sessions 6 months apart displayed excellent agreement over time. Furthermore, the PSSS scoring system correlated with another similar scoring system, the OSBD-r, confirming the criterion validity of the PSSS. Finally, construct validity was demonstrated by finding that the change in PSSS score happened consistently over time with respect to the initiation of the procedure and with a return to the presedation score after completion of the procedure.

We previously published the development and validation of the DOCS for use with video analysis of sedation strategies. In contrast to the DOCS, the PSSS is a much more clinically usable scale in that it requires evaluation in just 1 linear dimension rather than 4 separate axes and summation of scores in each axis. Although not as detailed as the DOCS, we believe the PSSS could be used to document clinical care on a daily basis. It could also be employed to evaluate the quality and safety of sedation techniques and compare the ability of a given technique to meet the requirements for a procedure. The scale could also be used to quantify the time required to attain acceptable procedural sedation conditions at the beginning of a procedure or the time required to come back to the baseline state before discharge. This type of detail is critical in comparing how well sedation techniques work and will yield much more useful information than the studies that currently only give the crudest data in terms of procedure completion and the presence of adverse events.

Similar to the DOCS, the PSSS could also be helpful in assessing alternative sedation techniques. Pure sedation scales do not allow comparison of the procedural conditions produced through the use of pharmacology with those accomplished through distraction or other noninvasive techniques. PSSS scoring allows comparison of a patient’s state during pediatric procedures regardless of the depth of sedation. For instance, a lumbar puncture performed with excellent distraction methods could be compared directly with one performed under deep sedation. At the same time, the scale would indicate potentially dangerous conditions present during the procedure and whether those states are due to lack of adequate sedation (and dangerous movement) or excessive sedation leading to airway compromise. Rather than promoting pharmacologic sedation, we hope this type of direct and quantifiable comparison will lead to more consideration of nonnonsedation interventions and promote the most effective strategies.

In developing the PSSS, we employed a human factors method with control theory and sought input from a variety of pediatric sedation providers. We believe the application of these concepts is helpful in evaluating an inherently complicated area of practice such as pediatric sedation. At the same time, we recognize the methodological limitations our study presents, including the fact that our scale attempts to measure a relatively new outcome variable, that of the state as opposed to the previously well-described sedation level. We also recognize the scale was validated in 1 institution with a particular set of sedation locations. Testing of this scale in a variety of settings is needed. Another limitation of our study is that the scale was not tested or validated using physicians representing all pediatric subspecialties that provide sedation care. On the other hand, we have shown excellent agreement across a wide variety of individuals who either provide sedation or help to intake and recover these patients, including pediatric anesthesiologists, dentists, trainees, Certified Registered Nurse Anesthetists, nurses, and medical assistants. Validation among other
pediatric professionals is ultimately needed, but we feel confident, given its consistency among our testing cohort, that the measure will be similarly consistent in its outcomes among pediatric subspecialties. We understand that the criterion and construct validity portions of the study were performed on patients receiving a relatively minimal sedation for a minimally painful intervention. Although the intra- and interrater reliability was excellent in these patients, regardless of level of sedation, future validation in construct and criterion dimensions should take place in patients receiving deep sedation levels for painful procedures.

Finally, we recognize that although we have shown the validity of the PSSS in several dimensions, we have not shown that adherence to a particular PSSS rating results in improved outcomes for patients or the proceduralists involved in these cases. Although we appreciate this, we would suggest there is “face validity” to the idea that a patient who is relaxed and not moving during a procedure is inherently in a better state for that procedure (as long as he or she is safe) when compared with a patient who is moving randomly and crying out during the procedure.

CONCLUSIONS
We have described a new scale specifically designed for evaluating pediatric patients undergoing sedation for diagnostic and therapeutic procedures. The PSSS measures the state of the patient during a procedure and (within the limits of this study) has been shown to be a consistent and valid measure. This scale could be used to help measure the effectiveness and efficiency of sedation practice with greater detail and precision than has been possible with previous sedation-oriented scales.

ACKNOWLEDGMENTS
The authors would like to thank all of the members of the expert panel for their time and effort in contributing to the components of the PSSS scale. In addition, we would like to thank Maria Ieni, MD, Daniel B. Fenster, MD, MS, John Babineau, MD, and Peter Dayan, MD, MSc, for their help in completing the criterion and construct validity portions of the study.

ABBREVIATIONS
CI: confidence interval
DOCS: Dartmouth Operative Conditions Scale
ICC: intraclass correlation coefficient
OSBD-r: Observational Scale of Behavioral Distress–revised
PSSS: Pediatric Sedation State Scale

REFERENCES


Validation of the Pediatric Sedation State Scale
Joseph P. Cravero, Nissa Askins, Patcharee Sriswasdi, Daniel S. Tsze, David Zurakowski and Sean Sinnott
Pediatrics 2017;139;
DOI: 10.1542/peds.2016-2897 originally published online April 28, 2017;
Validation of the Pediatric Sedation State Scale
Joseph P. Cravero, Nissa Askins, Patcharee Sriswasdi, Daniel S. Tsze, David Zurakowski and Sean Sinnott
Pediatrics 2017;139;
DOI: 10.1542/peds.2016-2897 originally published online April 28, 2017;

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/139/5/e20162897