Development of the Japanese version of the State Behavioral Scale for critically ill children

Haruhiko Hoshino,1,2 Hideaki Sakuramoto,3 Yuiiro Matsuishi,1,2 Nobutake Shimojo,1 Yuki Enomoto,1,4 Tatsuyuki Ohto,4 Takahiro Kido,4 Akira Ouchi,1,3 Masahiko Sumitani,5 Satoru Kawano,1 and Yoshiaki Inoue1

1Department of Emergency and Critical Care Medicine, Faculty of Medicine, University of Tsukuba, 2Pediatric intensive care unit, 3Intensive care unit, 4Department of Child Health, University of Tsukuba Hospital, Tsukuba, Ibaraki, Japan, and 5Department of Pain and Palliative Medicine, The University of Tokyo Hospital, Tokyo, Japan

Aims: The State Behavioral Scale (SBS) was developed to assess sedation states, including agitation, in pediatric patients on mechanical ventilation. The purpose of this study was to determine the reliability and validity of a back-translated Japanese version of the SBS.

Methods: Translation was done by the back-translation method followed by a prospective study in a Japanese intensive care unit. For reliability, a nurse/researcher pair evaluated SBS along eight dimensions (respiratory drive, response to ventilation, coughing, best response to stimulation, attentiveness to care provider, tolerance to care, consolability, and movement after consoled). For validity, SBS scores were compared to the Richmond Agitation–Sedation Scale and a visual analog scale (VAS).

Results: The original author approved the back-translated SBS. Thirty-one patients aged 0 weeks to 8 years were evaluated from 59 total critical pediatric patient encounters. The researcher and nurse SBS scores demonstrated excellent inter-rater reliability (weighted \( \kappa = 0.96, 95\% \text{ CI } 0.92–0.99 \)). In addition, there was a very strong correlation between the researcher and nurse VAS scores \( (r = 0.80, P < 0.001) \). Weighted kappa coefficients for the eight dimensions ranged from 0.71 (95% confidence interval, 0.55–0.88; consolability) to 0.89 (95% confidence interval, 0.80–0.98; best response to stimulation). In validity testing, nurse SBS and nurse VAS scores were strongly correlated \( (r = 0.80, P < 0.001) \) with the researcher SBS and researcher Richmond Agitation–Sedation Scale scores \( (r = 0.91, P < 0.001) \).

Conclusion: This study suggests that our Japanese version of the SBS is valid and reliable for evaluating sedation for critically ill children.

Key words: Agitation, pediatric intensive care, sedation, State Behavioral Scale

BACKGROUND

Critically ill pediatric patients may experience physical and psychological distress associated with physiological imbalances, environmental stressors, and treatments (e.g., mechanical ventilation) in the pediatric intensive care unit (ICU). Sedatives are often required to ameliorate such distress and associated anxiety as well as decrease oxygen consumption. However, excessive sedation causes extended ventilator usage that can precipitate iatrogenic withdrawal syndrome and delirium.1–3 It is therefore important to maintain an optimal sedation state.

Clinical recommendations for sedation consist of the COMFORT scale, COMFORT behavior scale, and State Behavioral Scale (SBS).4 Both the COMFORT and COMFORT behavior scales assess not only sedation state but also pain,5–7 precluding their use for precisely calculating sedation states. As treatment differs based on the sedation and pain states of the patient, they must be evaluated independently. To this end, the SBS measures only sedation states, including agitation states, in critically ill children on mechanical ventilation receiving sedatives (Table 1).8 In addition, the SBS serves as a useful screening tool for iatrogenic withdrawal symptoms and delirium in the pediatric ICU.9,10

However, a reliable and validated Japanese translation of the SBS has never been developed. The purpose of this study was to translate the original SBS into Japanese and...
evaluate the validity and reliability of this Japanese version of SBS.

METHODS

Back-translation

After receiving written permission from the original author, Dr. Martha A. Q. Curley, we began the SBS translation with the back-translation method to ensure linguistic and cultural equivalence. For this purpose, the adaptation of Brislin’s translation model by Jones and colleagues was specifically selected for this project as this method involves multiple people in the translation process.10

The translation process followed a series of steps (Fig. 1). In the first step, the main researcher translated the English version of SBS to Japanese to create a tentative translation. This was then back-translated by another bilingual expert to ensure linguistic and cultural equivalence. The adaptation of Brislin’s translation model by Jones and colleagues was specifically selected for this project as this method involves multiple people in the translation process.10

Table 1. Original version of the State Behavioral Scale

| Number | Description                                      |
|--------|--------------------------------------------------|
| -3     | Unresponsive                                     |
| 2      | Responsive to noxious stimuli                     |
| 1      | Responsive to gentle touch or voice              |
| 0      | Awake and able to calm                           |
| 1      | Restless and difficult to calm                   |
| 2      | Agitated                                         |

© 2018 The Authors. Acute Medicine & Surgery published by John Wiley & Sons Australia, Ltd on behalf of Japanese Association for Acute Medicine
Japanese version. The second step was for a professional translator not familiar with the SBS to retranslate this tentative Japanese version into a tentative English version. In the third step, a committee of eight medical professionals (two clinical researchers, two intensivists, two pediatricians, and two pediatric intensive care nurses) discussed the differences between the original English version and the tentative English version to reduce variability through the expertise of a multidisciplinary medical team. Based on the findings of this discussion, the Japanese version was revised to create the final Japanese version. In the fourth step, the final Japanese version of SBS was once again back-translated to English by a professional translator. In the final step, the final back-translation was then checked by the original author.

© 2018 The Authors. *Acute Medicine & Surgery* published by John Wiley & Sons Australia, Ltd on behalf of Japanese Association for Acute Medicine
Every effort was made to carefully execute all the steps in the translation processes to avoid the loss of any original meaning due to cultural or linguistic differences.

**Patient sampling**

The validity and reliability of our Japanese version of SBS was evaluated at the University of Tsukuba Hospital ICU (Tsukuba, Japan) between December 2010 and May 2011. Patients under 18 years old who stayed in the Tsukuba University Hospital ICU for more than 24 h were recruited. As a result, a total of 31 enrolled patients aged 0 weeks to 8 years were then evaluated from a total of 59 encounters. Exclusion criteria were a neuromuscular blockade, non-ventilated patients, and/or central nervous system disorders. This study was approved by the ethics board of the University of Tsukuba Hospital Clinical Research Ethics Review Committee (H22-623).

A pair consisting of a trained researcher (registered nurse) and a trained nurse simultaneously and independently assessed a convenience sample of pediatric ICU patients on the SBS (the training period was for 1 month), visual analog scale (VAS), and Richmond Agitation-Sedation Scale (RASS). We started data collection when the trained researcher and trained nurse agreed that they had reached a common understanding of the SBS, eight dimensions of SBS, RASS, and VAS. The SBS sedation depth (−3 to +2) and eight dimensions (respiratory drive, response to ventilation, coughing, best response to stimulation, attentiveness to care provider, tolerance to care, consolability, and movement after consoled) were evaluated. Each dimension contained three to six scores that were used to calculate sedation and agitation levels. The SBS was assessed in the same way as the previous study. The VAS and RASS were used to assess the criterion validity of the SBS. The trained researcher and trained nurse scored a VAS from 0 to 100 mm by marking a single line on a scale anchored by the terms unresponsive and combative for each patient (the center of the line indicated alert and calm). The VAS was assessed before the SBS and RASS to preclude any influence that the sedation scales could have had on the researcher’s and nurse’s assessment of the VAS. The RASS evaluates sedation and agitation states in critically ill adult patients with a 10 step scale (+4 combative to −5 unarousable; Table 2) and we used a Japanese version of RASS previously reported. The validity and reliability of RASS has already been established for children in pediatric ICUs.

Additional demographic and clinical data were recorded, including age, sex, Pediatric Risk of Mortality III score on admission, Pediatric Cerebral Performance Category, Pediatric Overall Performance Category, use of mechanical ventilation and sedative use at the time of assessment.

**Statistical analysis**

Patient characteristics analyzed were age, sex, cerebral and overall performance, mortality risk, use of mechanical ventilation, sedation, and analgesics. Inter-rater reliability was determined for researcher and nurse SBS (SBS and eight dimensions) evaluations using weighted kappa (quadratic weights).

For criterion validity testing, Spearman’s correlation coefficient was computed for the associations between nurse SBS scores and nurse VAS scores. For these calculations, SBS along eight dimensions and nurse VAS scores were used. To test the construct validity, researcher SBS scores and researcher RASS scores were compared using Spearman’s correlation coefficient. The SBS value comparisons were calculated with segregation into subgroups based on

---

**Table 2. Original version of the Richmond Agitation–Sedation Scale**

| Score | Term           | Description                                                                 |
|-------|----------------|------------------------------------------------------------------------------|
| 4     | Combative      | Overtly combative or violent; immediate danger to staff                      |
| 3     | Very agitated  | Pulls on or removes tube(s) or catheter(s) or has aggressive behavior toward staff |
| 2     | Agitated       | Frequent non-purposeful movement or patient—ventilator dyssynchrony         |
| 1     | Restless       | Anxious or apprehensive but movements not aggressive or vigorous             |
| 0     | Alert and calm | Not fully alert, but has sustained (more than 10 s) awakening, with eye contact, to voice |
| −1    | Drowsy         | Briefly (<10 s) awakens with eye contact to voice                            |
| −2    | Light sedation | Any movement (but no eye contact) to voice                                   |
| −3    | Moderate sedation | No response to voice, but any movement to physical stimulation             |
| −4    | Deep sedation  | No response to voice or physical stimulation                                |
| −5    | Unarousable    |                                                                              |
age and administration of midazolam. Comparisons were made with the Mann–Whitney U-test and Kruskal–Wallis rank-sum test/ANOVA, as appropriate.

All validity and reliability testing compared only unique first encounters to eliminate same-patient encounter bias. All statistical analyses were carried out using SPSS version 24 (SPSS Inc., Chicago, IL, USA). Data collection and analyses were independently undertaken to the fullest extent possible.

RESULTS

Back-translation

The final English version of SBS created by the back-translation method was submitted to the original author, who then approved the back-translated final English version of the SBS. The Japanese version of the SBS is displayed in Table S1.

Patient characteristics

During the test period, a total of 59 critical pediatric patient encounters in 31 patients were evaluated. The characteristics of the 59 ICU patient encounters are detailed in Table 3. The patient’s age ranged from 0 weeks to 8 years. Age groups were: 0–1 year, 29 (49%); 1–3 years, 12 (20%); 3–8 years, 18 (31%). For the SBS, 75% of the scores had a sedation range of –3 to –1, 17% scored 0, and 8% were in the agitation range (1–2). There was no significant difference with respect to age groups in the three scores (SBS, RASS, and VAS).

Reliability

In 59 paired observations by the researcher and nurse, both SBS scores showed excellent inter-rater reliability with a weighted kappa of 0.96 (95% confidence interval [CI], 0.92–0.99) and both VAS scores had excellent correlation (ρ = 0.80 [P < 0.001]). When analyzing only the first 31 encounters, the agreement was the same (κ = 0.95; 95% CI, 0.91–1.0; ρ = 0.74 [P < 0.001]). Calculated SBS agreements for each age group were κ = 0.92 (95% CI, 0.85–0.99) for ages 0–1 year, κ = 1.00 (95% CI, 1.0–1.0) for ages 1–3 years, and κ = 0.94 (95% CI, 0.86–1.0) for ages 3–8 years (all calculations had P < 0.001).

Weighted kappa coefficients for the eight dimensions of the SBS ranged from 0.71 (95% CI, 0.55–0.88; consolability) to 0.89 (95% CI, 0.80–0.98; best response to stimulation). When controlled for only the first observations (n = 31), weighted kappa coefficients ranged from 0.56 (95% CI, 0.28–0.84; movement after consoled) to 0.86 (95% CI, 0.67–1.0; consolability) as seen in Table 4.

Validity

A total of 59 encounters were tested for criterion validity. The results showed that nurse SBS scores were strongly correlated with nurse VAS scores (ρ = 0.80, P < 0.001; Fig. 2). These coefficients, representing eight dimensions of the nurse SBS and nurse VAS scores, were as follows: respiratory drive, ρ = 0.67 (P < 0.001); response to ventilation, ρ = 0.59 (P < 0.001); coughing, ρ = 0.72 (P < 0.001); best response to stimulation, ρ = 0.84; attentiveness to care provider, ρ = 0.71 (P < 0.001); tolerance to care, ρ = 0.74 (P < 0.001); consolability, ρ = 0.75 (P < 0.001); and movement after consoled, ρ = 0.57 (P < 0.001). Age group correlation between nurse SBS and VAS scores was as follows: 0–1 year, ρ = 0.89 (P < 0.001); 1–3 years, ρ = 0.84 (P < 0.001); and 3–8 years, ρ = 0.66 (P = 0.002). When controlled for only the first patient observations, the nurse

Table 3. Characteristics of the first encounters (n = 31) and all encounters (n = 59) of critical pediatric patients enrolled in this study

| Characteristic (first encounters) | n = 31 |
|-----------------------------------|--------|
| Age at enrollment (months), median (IQR) | 10 (2–42) |
| Sex, frequency (%) |        |
| Female | 15 (48) |
| Male | 16 (52) |
| Pediatric cerebral performance category median (IQR) | 1 (1–3) |
| Pediatric overall performance category median (IQR) | 1 (1–3) |
| Pediatric risk of mortality III scores (day 1), median (IQR) | 12 (7.25–20.8) |

IQR, interquartile range.
SBS scores correlated highly with the nurse VAS scores ($p = 0.74$, $P < 0.001$).

Two results verified construct validity. First, the researcher SBS scores were significantly associated with the researcher RASS scores ($p = 0.91$, $P < 0.001$; Fig. 3). Second, the SBS scores were significantly lower in patients treated with midazolam ($p = 0.03$). Age group correlation between researcher SBS and RASS scores were as follows: 0–1 year, $p = 0.86$ ($P < 0.001$); 1–3 years, $p = 0.88$ ($P < 0.001$); and 3–8 years, $p = 0.96$ ($P < 0.001$). When controlled for only the unique patient encounters ($n = 31$), there was a correlation between researcher SBS and RASS scores ($p = 0.90$, $P < 0.001$).

**DISCUSSION**

We constructed and clinically validated a Japanese translation of the SBS that can be used as a valid sedation scale for critically ill children. Accuracy was guaranteed not by word-for-word translation but rather by using the back-translation method. In order to take into account any cultural differences, we selected to use the Jones back-translation method, which involves multiple medical disciplines and cultures to reduce such differences.

There was a very strong inter-rater agreement rate between researcher SBS and nurse SBS scores. In addition, researcher VAS scores correlated highly with nurse VAS scores, leading to the conclusion that the SBS reflects both nurse and researcher opinions. Moderate to excellent inter-rater reliability was indicated in all eight SBS dimensions. This kappa scores was more than to the original study. The previous study has already described the reliability of the modified Japanese version of SBS, however, it did not use the back-translation method, agreement was only moderate, and dimension of “attentiveness to care provider” was poor.
The difference in sample size between the present and previous studies could cause such a difference in outcomes, therefore, more future studies on the validity and reliability of SBS could be required.

Our Japanese version of the SBS was determined to have high validity as eight dimensions of the nurse SBS and nurse VAS scores showed moderate to strong correlation. Although there is no gold standard for comparison of sedation scales, VAS has been used in similar studies to validate sedation scales, and for this reason we selected VAS scores as our standard. We used RASS as a sedation scale to measure validity as it has been validated in pediatric intensive care patients. Therefore, RASS has been tested and established as a valid measure of the clinical evaluative powers of nurses with regards to sedation states. Interestingly, validity and reliability results did not change in any age group when controlled for only the unique patient encounters (n = 31), further bolstering the efficacy of the translated SBS. Taken together, our results lead us to the conclusion that our back-translated SBS is the first Japanese version that has been tested for validity and reliability and will therefore be useful in evaluating pediatric sedation and agitation states within Japanese pediatric patients.

Limitations

Several limitations of this study should be acknowledged. First, our sampling methodology relied on patient availability, possibly imparting selection bias. Second, our study did not calculate exact patient sample size, although it was still similar to previous studies used to evaluate sedation tools. Third, we analyzed the assessment from only one paired researcher and nurse. In the original study, the assessment nurse pair was randomly chosen from five trained nurses to assess the SBS. Thus, it was possible that the pair unconsciously gave VAS scores that were conveniently consistent with yet-to-be-determined SBS and RASS. Fourth, non-ventilated patients were excluded from SBS assessment, although there is a crucial need for sedation assessment in this population. For this reason, it might be necessary to develop a modified SBS adapted for non-ventilated patients. Finally, this study was undertaken at only a single ICU as an initial validation test. Future studies at multiple ICUs will provide foundational data for testing of this new Japanese translation.

CONCLUSION

This study suggests that our Japanese version of the SBS is valid and reliable for evaluating sedation in pediatric intensive care patients.

ACKNOWLEDGEMENTS

We are grateful to Dr. Martha A.Q. Curley for granting us permission to translate SBS and providing guidance. We would also like to thank Dr. Bryan J. Mathis of the University of Tsukuba Medical English Communication Center for critical editing of our manuscript.

DISCLOSURE

Approval of the research protocol: The present study was approved by the ethics board of University of Tsukuba Hospital.

Informed consent: N/A.

Registry and registration no. of the study/trial: UMIN 000004713. Registered 14 December 2010.

Animal studies: N/A.

Conflict of interest: None declared.

REFERENCES

1 Smith HAB, Gangopadhyay M, Goben CM et al. Delirium and benzodiazepines associated with prolonged ICU stay in critically ill infants and young children. Crit. Care Med. 2017; 45: 1427–35.
2 Franck LS, Scoppettuolo LA, Wypij D, Curley MAQ. Validity and generalizability of the withdrawal assessment tool-1 (WAT-1) for monitoring iatrogenic withdrawal syndrome in pediatric patients. Pain 2012; 153: 142–8.
3 Traube C, Silver G, Kearney J et al. Cornell assessment of pediatric delirium: a valid, rapid, observational tool for screening delirium in the PICU*. Crit. Care Med. 2014; 42: 656–63.
4 Harris J, Ramelet AS, van Dijk M et al. Clinical recommendations for pain, sedation, withdrawal and delirium assessment in critically ill infants and children: an ESPNIC position statement for healthcare professionals. Intensive Care Med. 2016; 42: 972–86.
5 Van Dijk M, De Boer JB, Koot HM, Tibboel D, Passchier J, Duivenvoorden HJ. The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. Pain 2000; 84: 367–77.
6 Ista E, van Dijk M, Tibboel D, de Hoog M. Assessment of sedation levels in pediatric intensive care patients can be improved by using the COMFORT “behavior” scale. Pediatr. Crit. Care Med. 2005; 6: 58–63.
7 Van Dijk M, Bouwmeester NJ, Duivenvoorden HJ et al. Efficacy of continuous versus intermittent morphine administration after major surgery in 0-3-year-old infants; a double-blind randomized controlled trial. Pain 2002; 98: 305–13.
8 Curley MAQ, Harris SK, Fraser KA, Rita A, Arnold JH. State behavioral scale (SBS) a sedation assessment instrument for infants and young children supported on mechanical ventilation. Pediatr. Crit. Care Med. 2006; 7: 107–14.
9 Franck LS, Harris SK, Soetenga DJ, Amling JK, Curley AQ. NIH Public Acces The withdrawal assessment tool – version 1 (WAT-1): an assessment instrument for monitoring opioid and benzodiazepine withdrawal symptoms in pediatric patients. Pediatr. Crit. Care Med. 2009; 9: 573–80.
10 Jones PS, Lee JW, Phillips LR, Zhang XE, Jaceldo KB. An adaptation of Brislin’s translation model for cross-cultural research. Nurs. Res. 2001; 50: 300–4.
11 Ely EW, Truman B, Thomason JWW et al. Monitoring sedation status over time in ICU patients reliability and validity of the Richmond agitation-sedation scale (RASS). JAMA 2003; 289: 2983–91.
12 Unoki T, Sakuramoto H, Okimura A et al. Richmond agitation-sedation scale development of a Japanese version of Richmond agitation-sedation scale. J. Jpn Soc. Intensive Care Med. 2010; 17: 73–4.
13 Kerson AG, Demaria R, Mauer E et al. Validity of the Richmond agitation-sedation scale (RASS) in critically ill children. J. Intensive Care 2016; 4: 65.
14 Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated pediatric risk of mortality score. Crit. Care Med. 1996; 24: 743–52.
15 Fiser DH. Assessing the outcome of pediatric intensive care. J. Pediatr. 1992; 121: 68–74.
16 Koizumi T, Yamamoto K, Usami Y et al. The reliability of a modified Japanese version of State Behavioral Scale for assessing the sedation status in critically ill pediatric patients. J. Jpn Soc. Intensive Care Med. 2015; 22: 189–93.
17 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: 1271–5.
18 Sessler CN, Gosnell MS, Grap MJ et al. The Richmond agitation-sedation scale. Am. J. Respir. Crit. Care Med. 2002; 166: 1338–44.
19 Malviya S, Voepel-Lewis T, Tait AR, Merkel S, Tremper K, Naughton N. Depth of sedation in children undergoing computed tomography: validity and reliability of the University of Michigan Sedation Scale (UMSS). Br. J. Anaesth. 2002; 88: 241–5.
20 Ambuel B, Hamlett KW, Marx CM, Blumer JL. Assessing distress in pediatric intensive care environments: the COMFORT scale. J. Pediatr. Psychol. 1992; 17: 95–109.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

Table S1. Japanese version of the State Behavioral Scale.