

# Development of an Instrument for a Primary Airway Provider's Performance With an ICU Multidisciplinary Team in Pediatric Respiratory Failure Using Simulation

Akira Nishisaki MD MSCE, Aaron J Donoghue MD MSCE, Shawn Colborn RRT-NPS, Christine Watson RN, Andrew Meyer MD, Dana Niles MSc CCRC, Ram Bishnoi MBBS, Roberta Hales RRT-NPS NRN, Larissa Hutchins RN, Mark A Helfaer MD, Calvin A Brown III MD, Ron M Walls MD, Vinay M Nadkarni MD, and John R Boulet PhD

**OBJECTIVE:** To develop a scoring system that can assess the multidisciplinary management of respiratory failure in a pediatric ICU. **METHODS:** In a single tertiary pediatric ICU we conducted a simulation-based evaluation in a patient care area auxiliary to the ICU. The subjects were pediatric and emergency medicine residents, nurses, and respiratory therapists who work in the pediatric ICU. A multidisciplinary focus group with experienced providers in pediatric ICU airway management and patient safety specialists was formed. A task-based scoring instrument was developed to evaluate a primary airway provider's performance through Healthcare Failure Mode and Effect Analysis. Reliability and validity of the instrument were evaluated using multidisciplinary simulation-based airway management training sessions. Each session was evaluated by 3 independent expert raters. A global assessment of the team performance and the previous experience in training were used to evaluate the validity of the instrument. **RESULTS:** The Just-in-Time Pediatric Airway Provider Performance Scale (JIT-PAPPS) version 3, with 34 task-based items (14 technical, 20 behavioral), was developed. Eighty-five teams led by resident airway providers were evaluated by 3 raters. The intraclass correlation coefficient for raters was 0.64. The JIT-PAPPS score correlated well with the global rating scale ( $r = 0.71, P < .001$ ). Mean total scores across the teams were positively associated with resident previous training participation ( $\beta$  coefficient  $7.1 \pm 0.9, P < .001$ ), suggesting good validity of the scale. **CONCLUSIONS:** A task-based scoring instrument for a primary airway provider's performance with a multidisciplinary pediatric ICU team on simulated pediatric respiratory failure was developed. Reliability and validity evaluation supports the developed scale. *Key words:* intubation; pediatric; ICU; simulation; training; residents; nurses; respiratory therapists; airway. [Respir Care 2012;57(7):1121–1128. © 2012 Daedalus Enterprises]

## Introduction

Tracheal intubation in the pediatric ICU (PICU) is a life-saving procedure for critically ill children. The major-

ity of the airway management procedures in the PICU are non-elective and urgent.<sup>1,2</sup> Skilled laryngoscopists and

---

Dr Nishisaki, Dr Donoghue, Ms Niles, Dr Bishnoi, Ms Hales, and Dr Nadkarni are affiliated with the Center for Simulation, Advanced Education, and Innovation; Drs Nishisaki, Donoghue, Helfaer, and Nadkarni are affiliated with the Department of Anesthesiology and Critical Care Medicine; Mr Colborn is affiliated with the Department of Respiratory Care; Ms Watson and Ms Hutchins are affiliated with the Department of Nursing; and Dr Meyer is affiliated with the Division of Neonatology, Department of Pediatrics, the Children's Hospital of Philadelphia, Philadelphia, Pennsylvania. Drs Brown and Walls are affiliated with the Department of Emergency Medicine, Brigham and Women's Hospital, Boston, Massachusetts. Dr Boulet is affiliated with the Foundation for Advancement of International Medical Education and Research, Philadelphia, Pennsylvania.

---

Drs Nishisaki, Walls, Niles, and Nadkarni were partly supported by grant HS016678–01 from the Agency for Healthcare Research and Quality. The authors have disclosed no conflicts of interest.

Supplementary material related to this paper is available at <http://www.rcjournal.com>.

Correspondence: Akira Nishisaki MD, Department of Anesthesiology and Critical Care Medicine, The Children's Hospital of Philadelphia, Main Building 8NE, Suite 8566, 34th Street and Civic Center Boulevard, Philadelphia PA 19104. E-mail: Nishisaki@email.chop.edu

DOI: 10.4187/respcare.01472

trained multidisciplinary teams are necessary for this critical procedure.<sup>3,4</sup> Unfortunately, this procedure is often troubled with unexpected adverse tracheal intubation associated events.<sup>1,2,4-9</sup> Those unwanted events are sometimes associated with providers' skills and experiences. Importantly, some of the required skills are not only technical psychomotor skills, rather, behavioral and teamwork skills similar to adult and pediatric resuscitation.<sup>3</sup> Therefore it is essential to have individual and team-based airway management training in the PICU for the safety of critically ill children.

Simulation-based individual and team training have been proven to be effective in improving technical skills and teamwork skills in central line insertion, trauma resuscitation, or advanced cardiac life support training in the clinical setting.<sup>10-12</sup>

To assess the effectiveness of the training in the simulation-setting or at the bedside, a reliable and valid measurement for provider performance is essential.<sup>13,14</sup> However, this type of scale for pediatric airway management has not been available. We therefore attempted to develop and evaluate an objective instrument to rate a primary airway provider's performance with a PICU multidisciplinary team in pediatric airway management. We sought to evaluate the properties of this instrument for its reliability and validity in simulation as the first step.

## Method

### Setting

This study was conducted in the PICU at a single tertiary children's hospital, with institutional review board approval. The PICU is a 45 bed, tertiary PICU staffed with 20 faculty, 15 pediatric critical care medicine fellows, and approximately 150 nurses. Eight to ten pediatric or emergency medicine (non-anesthesiology track) residents rotate in the PICU every month. At least 4 respiratory therapists are always on duty in the unit at any given time. Each ICU intubation is supervised by a pediatric critical care medicine board certified/eligible ICU attending and/or a PICU fellow. Other members of the bedside airway team include a resident, 2 ICU nurses, and at least one ICU respiratory therapist.

### Design

This study was conducted as part of a quality improvement educational study: Just-in-Time Simulation Training on Tracheal Intubation Procedure. This educational intervention focused on the basic and advanced airway management by a non-anesthesiology (pediatric or emergency medicine) resident trainee, PICU nurse, and a respiratory therapist. This study design was published elsewhere.<sup>4</sup>

## QUICK LOOK

### Current knowledge

Pediatric airway management in the pediatric ICU is a team based activity, and the risks and the skills required for successful airway management are different from airway management in the operating room. Team based airway management training in the pediatric ICU may enhance safety.

### What this paper contributes to our knowledge

The Just-in-Time Pediatric Airway Provider Performance Scale (JIT-PAPPS) was developed to assess a physician airway leader's ability to stabilize a patient with impending respiratory failure. Healthcare Failure Mode and Effect Analysis was used to develop the scale for a successful airway management of pediatric respiratory failure by a team in the pediatric ICU. The JIT-PAPPS was found to be valid and reliable.

Briefly, the airway team consisted of one of the on-call residents, PICU nurses, and respiratory therapists, who received a 20-min multidisciplinary simulation-based airway management training in the morning. The on-call resident also received a 10-min skill refresher training for the psychomotor intubation skill, using a pediatric airway trainer (Laerdal, Wappinger Falls, New York). One confederate was present to act a role for the second PICU nurse, who prepared the medication and served as a documenter, as requested by the team. When the PICU nurse was not available for participation, another confederate was used as a PICU role nurse.

Two scenarios were used, in a randomized fashion, for the simulation training (see Appendix A in the supplementary materials at <http://www.rcjournal.com>). Both scenarios involved the management and placement of a tracheal tube in the PICU. Each scenario started with a different stem (scenario A was bronchiolitis, scenario B was pneumonia). The initial and programmed vital sign changes for the provided actions were identical. Each scenario required the same skills to resuscitate the patient and perform intubation. Each scenario was divided into 2 stages, based on the required skills: basic airway management stage (requirement of proper bag-valve-mask ventilation), and advanced airway management stage (requirement of tracheal intubation and appropriate confirmation).

### Development of the Trichotomous Just-in-Time Pediatric Airway Performance Scale

The Just-in-Time Pediatric Airway Provider Performance Scale (JIT-PAPPS) was developed de novo, be-

cause similar scales were not available at the time of the study. This scale was developed to assess a physician airway leader's ability to stabilize a patient with impending respiratory failure, without an anticipated difficult airway. Healthcare Failure Mode and Effect Analysis (HFMEA) was used to develop the scale for successful airway management of pediatric respiratory failure by a team in the PICU.<sup>15,16</sup>

Initially, a multidisciplinary focus group was formed, with a critical care attending physician, a respiratory therapist with more than 10 years of PICU experience, an advanced practice nurse in critical care, and a patient safety specialist with experience with HFMEA.

Next, the focus group analyzed the typical steps to rescue an infant with acute respiratory failure. Four critical processes were identified: to recognize impending/existing respiratory failure; to provide effective bag and mask ventilation; to prepare for orotracheal intubation; and to perform orotracheal intubation successfully. For the purpose of the score development, the 2–4 processes were chosen for further analysis. The team then identified subprocesses (ie, the steps within the process). Next, for each subprocess the group assigned 3 numeric values, using a probabilistic risk assessment model: the likelihood of occurrence (how likely is it that this failure mode will occur?); likelihood of detection (if this failure mode occurs, how likely is it that the failure will be detected?); and severity (if this failure mode occurs, how likely is it that harm will occur?).<sup>17</sup> Each value ranged from 1–10 (1 = very unlikely, 10 = highly likely). Risk propriety score was calculated as a product of those 3 numeric values.

Then the group provided a point for each subprocess, based on this risk priority score. For each 100 risk priority score, each subprocess was given one additional point. One point was given to each subprocess with less than a 100 risk priority score.

Each observable subprocess was then converted into an item on a task-based evaluation list, as a preliminary version of the JIT-PAPPS. By consensus, each item was categorized either in the technical domain or the non-technical (behavioral) domain, based on the teamwork conceptual model by Fletcher et al.<sup>18</sup>

Using an expert research meeting with a video clip of the simulation, we refined the scales, and additional items were added. After the preliminary use we revised the scale from a dichotomous system to a trichotomous system (0 = not done, 1 = partially done, done incorrectly, and 2 = done correctly), to improve usability,<sup>19–22</sup> and, finally, JIT-PAPPS was developed. Both domain scores are calculated as the weighted sum of items that belong to each domain. A total score is calculated as a sum of technical and non-technical (behavioral) domain score. In addition, a global assessment score was collected for basic airway management (from recognition of impending respiratory failure to

the decision to proceed to advanced airway management) and for advanced airway management (pronunciation of advanced airway management plan to the completion of tracheal intubation with appropriate confirmation). Those were used as a reference to validate itemized JIT-PAPPS scores. Operational definitions were also developed through this process (see Appendix B in the supplementary materials at <http://www.rcjournal.com>).

### JIT-PAPPS Video Review and Data Collection

Each simulation session was videotaped with 3 video cameras with different angles. One combined file was generated, with 4 windows, with 3 video images and a simulator monitor display. This file also included automated logs from a simulator and a sound recording (AVS, Laerdal, Wappinger Falls, New York). One facilitator at the training immediately rated the performance, using the JIT-PAPPS on site. Two independent raters blinded to the participant training levels rated the performance, using the video file, in a randomized order. Those raters were both content and simulation experts in pediatric airway management.

### Evaluation of the Property of JIT-PAPPS and Statistical Analyses

The properties of the JIT-PAPPS were evaluated based on a fully crossed design: 3 raters evaluated all teams. For all items on the JIT-PAPPS, detailed psychometric analyses were performed. Item difficulty—the mean score of the team over the 3 raters (range 0–2)—was reported. A value of 2 would indicate that all teams received credit for the item by all 3 raters. The second measure was item discrimination, which is the correlation between the item-level score and the total JIT-PAPPS score. Here, higher values indicate that the item is able to discriminate low- and high-ability teams. In some instances (ie, all or no teams received credits), the discrimination cannot be calculated. The third measure was inter-rater correlation. This was estimated from the variant component analysis, using a fully crossed design with 3 raters. This is equivalent to the correlation estimated for randomly selected 2 raters.

A similar analysis strategy was employed for the domain and total scores. Here we evaluated the correlation between the scores of each domain (technical and non-technical) and total scores.

A team by rater generalizability study was conducted to partition the sources of variability in the JIT-PAPPS technical, non-technical (behavioral), and total scores. The estimated variance component analysis was performed to calculate an overall measure of inter-rater reliability. This represents the average correlation in scores between any 2 randomly selected raters. To evaluate the validity of the

## INSTRUMENT FOR A PRIMARY AIRWAY PROVIDER'S PERFORMANCE

score, 2 strategies were adopted. First we evaluated the JIT-PAPPS scores against the global (holistic) rating. Here we calculated the correlation between JIT-PAPPS scores and global rating for all, and for basic and advanced airway management separately. Second, linear regression analysis was applied with the resident previous Just-in-Time training participation as an independent variable and the JIT-PAPPS scores as dependent variables. Here the association of previous participation with improved score would demonstrate the effectiveness of the training, resulting in the validation of the JIT-PAPPS score.

Descriptive statistics with mean  $\pm$  SD were calculated. Correlation coefficient with significance is reported. Variance component analysis, using a generalizability theory for a fully crossed design, was used to assess each airway provider's performance.<sup>23</sup> This analysis was necessary because all performances were reviewed by 3 raters independently. Inter-rater reliability was calculated based on this analysis. Where appropriate, linear regression with 95% CI was reported and displayed in the figure. A 2-tailed test with  $P < .05$  was used as a cutoff. Statistics software (Stata 11.0, StataCorp, College Station, Texas) was used for analysis throughout.

### Results

#### Subjects

A total of 85 team performances were rated by the 3 raters. Thirty-two teams had a resident without previous Just-in-Time training, 27 teams had a resident with one previous Just-in-Time training experience, 19 had a resident with 2 previous trainings, and 7 had a resident with 3 previous trainings. Fifteen sessions did not have a bedside PICU nurse participation, and a confederate PICU nurse was used instead.

#### Just-in-Time Pediatric Airway Provider Performance Scale

A newly developed JIT-PAPPS version 3 contains a total of 34 items: 14 items in the technical domain and 20 items in the behavioral (non-technical) domain (Table 1). Each item has 1–6 points as a weight, based on the HFMEA. The maximum possible technical domain score is 62, the behavioral domain score is 80, and the total score maximum is 142. Items 1–10 represent basic airway management (possible maximal score = 34), and items 11–34 represent the advanced airway management (possible maximal score = 108).

Three raters (one on-site rater/facilitator and 2 blinded video raters) graded a total of 85 team performances individually. The mean total score was  $103.9 \pm 15.8$ , the technical domain score was  $50.8 \pm 8.0$ , and the behavioral

Table 1. Just-in-Time Pediatric Airway Provider Performance Scale, Version 3

Item	Domain	Points	Checkpoint (done = 2, partially done = 1, not done = 0)
1	Nontechnical	1	Identify oneself
2	Nontechnical	1	Call for help
3	Technical	1	Put gloves on hands
4	Technical	2	Open airway with head-tilt chin-lift or jaw thrust within <i>first</i> 15 s
5	Technical	1	Choose right size mask
6	Technical	2	Check O <sub>2</sub> source is turned on; if not, turn it on
7	Technical	1	Apply mask correctly
8	Technical	3	Provide bag and mask ventilation to see chest rise
9	Nontechnical	3	Ask for blood pressure measurement during bag and mask ventilation
10	Nontechnical	2	Decide to intubate within 60 s after bag and mask ventilation is started
11	Nontechnical	1	Notify the team for intubation
12	Nontechnical	4	Call for suction system
13	Nontechnical	3	Call for oral airway
14	Nontechnical	1	Call for endotracheal tube
15	Technical	2	<i>Correct size</i> of endotracheal tube is called
16	Nontechnical	1	Call for laryngoscope
17	Nontechnical	1	Call for colorimetric end-tidal CO <sub>2</sub> detector
18	Technical	2	Wear mask with eye protection
19	Nontechnical	2	Call for sedative and/or narcotic medication
20	Nontechnical	2	Call for paralytic medication
21	Nontechnical	1	Confirm intravenous access is functional
22	Nontechnical	2	Confirm team crew at specific task (intubation assistant, person to give medication, person who watches the monitor)
23	Nontechnical	3	Ask for blood pressure cycle measurement before induction
24	Nontechnical	3	Ask for cricoid pressure when sedative/narcotics are given (before paralytics)
25	Technical	2	Stop bag and mask ventilation at <i>correct timing</i> (after paralyzed) for intubation
26	Technical	1	Hold laryngoscope with left hand
27	Technical	6	Be able to visualize a vocal cord
28	Technical	4	Intubate in trachea
29	Technical	2	Primary confirmation of endotracheal intubation
30	Technical	2	Secondary confirmation of endotracheal intubation
31	Nontechnical	2	Holding endotracheal tube until it is secured
32	Nontechnical	2	Call for chest x-ray
33	Nontechnical	2	Confirm endotracheal tube placement by chest x-ray
34	Nontechnical	3	React to hypotension after intubation

domain score was  $53.1 \pm 10.8$  (Table 2). The mean basic airway management score was  $27.3 \pm 4.7$ , and the advanced airway management score was  $76.6 \pm 13.8$ .

Table 2. Just-in-Time Pediatric Airway Provider Performance Scale: Total and Domain Specific Scores\*

Domain	Possible Maximal Score	JIT-PAPPS Score (mean ± SD)	Correlation Coefficient (r) With Total JIT-PAPPS Score
Total	142	103.9 ± 15.8	1.00
Technical	62	50.8 ± 8.0	0.77
Behavioral	80	53.1 ± 10.8	0.88
Basic airway management	34	27.3 ± 4.7	0.53
Advanced airway management	108	76.6 ± 13.8	0.96

\* All *P* values < .01.

JIT-PAPPS = Just-in-Time Pediatric Airway Provider Performance Scale, version 3.

Table 3. Variance Component Analysis for the Just-in-Time Pediatric Airway Provider Performance Scale, Version 3

Variance Component	Percent*
Airway provider performance	64.0
Rater	12.6
Residual	23.4

\* Based on 255 observations (3 ratings per each session for 85 teams).  $R^2 = 0.84$ .

**Reliability Assessment**

Based on the estimated variance components and the 3 raters, inter-rater reliability was calculated as 0.64. The variance attributed to the rater was small (12.6% of total variance), indicating that the raters provided similar scores over different teams (Table 3).

**Validity Assessment**

**Comparison of JIT-PAPPS and Global (Holistic) Rating.** For each team performance, global assessment and the total scores on JIT-PAPPS version 3 were assigned by all 3 raters. The correlation coefficient for those 2 scale scores was  $r = 0.71$  ( $P < .001$ ), as shown in Figure 1.

For basic airway management the correlation between the JIT-PAPPS version 3 score (items 1–10) and global assessment for basic airway skills was moderate ( $r = 0.48$ ,  $P < .001$ ). For advanced airway management the correlation was moderate ( $r = 0.64$ ,  $P < .001$ ).

**Analysis of JIT-PAPPS Scores By Training Participation.** Mean total scores across the teams were positively significantly associated with resident previous training participation ( $\beta$  coefficient  $7.1 \pm 0.9$ , degree of freedom = 1,  $P < .001$ , Fig. 2), suggesting construct validity of the scale.

**Comparison of JIT-PAPPS Domain Scores to Total Scores.** The technical domain in JIT-PAPPS version 3 had good correlation with the JIT-PAPPS total score

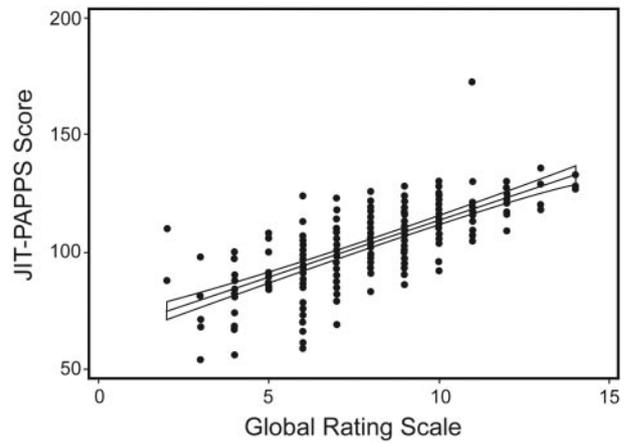


Fig. 1. Just-in-Time Pediatric Airway Provider Performance Scale score versus global rating scale score. The line displays the regression line with 95% confidence interval.

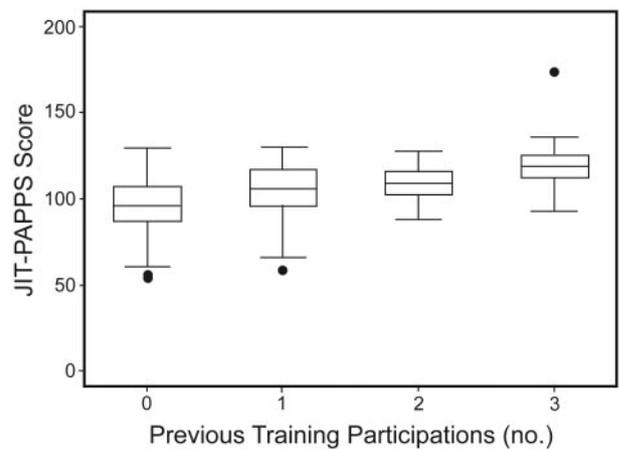


Fig. 2. Just-in-Time Pediatric Airway Provider Performance Scale score versus number of previous training participations. In the boxes, the middle line represents the median, the upper and lower lines of the box represent the 25th and 75th percentiles, the whisker bars represent the 5th and 95th percentiles, and the dots represent outliers (< 5th or > 95th percentile).

( $r = 0.77$ ,  $P < .001$ , see Table 2). The non-technical domain also had high correlation with the JIT-PAPPS total score ( $r = 0.88$ ,  $P < .001$ ). The technical score correlated modestly with the nontechnical (behavioral) score ( $r = 0.38$ ,  $P < .001$ ). The JIT-PAPPS scores in basic airway management (items 1–10) correlated modestly with the total scores ( $r = 0.53$ ,  $P < .001$ ). The JIT-PAPPS scores in advanced airway management (items 11–34) correlated highly with the total scores ( $r = 0.96$ ,  $P < .001$ ).

**Item-Level Analysis.** Item-level analyses demonstrated that, overall, the tasks were not difficult (mean score = 1.5 for all items, Table 4). The items “call for airway” and “ask for cricoid pressure” turned out to be difficult items (mean score < 0.5). “Wear mask with eye protection” was also difficult (mean = 0.4), revealing the discrepancy

Table 4. Item-Level Analysis: Difficulty, Discrimination, Inter-rater Correlation

Item	Difficulty* (mean raw score: 0–2)	Discrimination†	Inter-rater Correlation‡
1	0.7	0.36	0.01
2	1.9	0.14	0.11
3	1.4	0.30	0.66
4	1.3	0.22	0.10
5	2.0	-0.08	0.00
6	1.8	0.09	0.03
7	2.0	0.03	0.00
8	1.8	0.35	0.30
9	1.5	0.32	0.57
10	1.6	0.40	0.22
11	1.8	0.25	0.00
12	1.4	0.30	0.35
13	0.3	0.23	0.28
14	1.9	0.03	0.07
15	1.9	0.19	0.03
16	1.8	0.38	0.06
17	1.2	0.42	0.08
18	0.4	0.35	0.55
19	2.0	0.02	0.19
20	1.9	0.22	0.40
21	0.8	0.38	0.15
22	1.2	0.48	0.05
23	1.2	0.48	0.49
24	0.1	0.21	0.03
25	1.5	0.50	0.32
26	2.0	NA	NA
27	1.8	0.21	0.23
28	1.4	0.34	0.74
29	1.9	0.40	0.64
30	1.9	0.46	0.50
31	1.9	0.45	0.66
32	1.8	0.47	0.80
33	1.7	0.44	0.69
34	1.3	0.49	0.80

\* The mean of the entire raw score was 1.5 (raw score range 0–2).  
 † The discrimination is the correlation between the item-level score and the total checklist raw score. A higher value (> 0.30) indicates that the item was able to discriminate between low- and high-ability teams. In some instances (ie, all or no teams received the same scores), the discrimination value cannot be calculated.  
 ‡ The inter-rater correlation is the correlation with any of 2 rater's scores. This was calculated based on a G-study.  
 NA = not applicable

between the current bedside practice and the ideal practice described by the focus group. Discrimination statistics indicated that overall skills in tracheal intubation (items 25–34) had higher discriminative ability, compared to skills associated with basic airway management (items 1–10). This indicates that many teams performed equally in the basic airway management, while the teams performed differently based on their skills in the advanced airway management. The item-level inter-rater correlation statistics from a G study showed that the items in advanced airway

management had higher inter-rater correlation in general.<sup>23</sup> In some items it is more difficult to evaluate the physician airway provider's performance. For example, "Notify the team for intubation" had a correlation of 0.00, indicating that the non-verbal communication among team members was interpreted differently by raters. A similar finding was observed for the item "Call for laryngoscope." The laryngoscope was typically prepared by a respiratory therapist without being requested explicitly by a resident or a bedside nurse. This phenomenon caused raters to give different points to the airway provider's performance.

**Discussion**

We attempted to develop a usable and reliable instrument to assess the airway provider's performance with a multidisciplinary team for pediatric airway management. We utilized 2 approaches, a trichotomous task-based evaluation approach, and HFMEA by an expert focus group, to enhance content validity. The trichotomous task-based assessment was chosen because of its wide acceptance by clinical educators.<sup>19-22</sup> We adopted a typical clinical scenario with pediatric acute respiratory failure.

HFMEA with risk priority score was adopted because the investigators shared an underlying belief: all tasks are not equally important to achieve a successful and safe pediatric airway management. To improve the objectivity on those item-specific weights, a risk priority score based on HFMEA analysis was used. Our goal for this task-based scale was to develop a clinically meaningful, tailored instrument for bedside clinical use in the PICU.

As a result, we were able to develop a feasible instrument to assess the primary airway provider's performance on the management of acute respiratory failure. The inter-rater reliability of this instrument is good (0.64), with small variance component (12.6%). Each resident airway provider's performance improved substantially when they returned to a simulation session, while their multidisciplinary team members were different every time.

We used both an on-site assessment and a video-based assessment with independent raters. The video-based assessment resulted in some difficulty in rating behavioral items. Some items were never verbalized or explicitly communicated but were executed appropriately. The scores were split when a primary resident airway provider did not explicitly state the correct action, but the multidisciplinary team members were efficient in preparing and executing the correct action. Sometimes this occurred because of non-verbal cues from a resident airway provider. Overall this was rated as a credit to the airway provider, as defined in the operational definitions (see Appendix B), but interpreted differently by the raters, based on the situation. This improved when we had trained confederates (actors) for the training with explicit scripts.

Our raters for this study were all at senior ICU fellow or attending physician level, with substantial knowledge in simulation-based training. This contributed to high inter-rater reliability. If the JIT-PAPPS is used by novices, appropriate training with a proof of good inter-rater reliability with experts is recommended.

The global (holistic) evaluation of the team performance correlated well with the JIT-PAPPS scores. This supports the validity of this newly developed scale.

The merit in having a task-based instrument for simulation-based pediatric airway management is substantial. First, we need to have a mechanism to assess our training effect based on a performance during simulation. Pediatric basic and advanced airway management for trainees have become quite limited over years, despite the fact that they remain required in the pediatric residency program.<sup>4,24-26</sup> In the era of patient safety we cannot have our trainees practice on our sick patients in neonatal and pediatric ICUs. Some may argue that training in operating suites may be a better curriculum to teach airway management skills. However, the airway management in the ICU requires different technical and teamwork skills, as demonstrated in our HFMEA and the JIT-PAPPS as the end product. Simulation-based multidisciplinary training is suitable for this purpose. In our training curriculum we provided a basic technical psychomotor training with a task trainer. This was done repeatedly before the team training started. With the JIT-PAPPS we were able to measure and document the airway provider's performance and his or her improvement over time.

Second, not limited to the simulation environment, we may be able to measure the airway provider's performance at the bedside with the JIT-PAPPS in the future.<sup>3</sup> This step is essential to evaluate the transference of the effectiveness of the simulation-based education.<sup>27,28</sup>

This study is not free from limitations, and therefore warrants readers' discretion for interpretation of the results. The specific weights assigned to each item came from HFMEA analysis with the focus group. The transformation to weights from each risk priority score is arbitrary. Although overall inter-rater reliability was good, several items for behavioral tasks had poor inter-rater correlation. We acknowledge that this is due to the difficulty in assessing those tasks using recorded video files. We also need to tighten the operational definitions of how the rating should be for the actions the team members supported when the primary airway provider did not explicitly state or perform the action. This is not simple, though, since non-verbal communication between the primary airway provider and the multidisciplinary team members might have occurred during the simulation. We also need to evaluate the performance of JIT-PAPPS version 3 on high-end providers such as PICU fellows and attending physicians, to evaluate its discriminative validity.

## Conclusions

In summary, we developed a task-based scale for pediatric basic and advanced airway management in the PICU. The scale demonstrated good usability and reliability. Validity of the scale was also demonstrated in several ways. Further refinement with clear operational definition of the scale and demonstration of construct validity at the bedside are necessary.

## ACKNOWLEDGMENTS

We thank all of pediatric residents, ICU nurses, and respiratory therapists who participated in or supported this study. We deeply thank Stephanie Tuttle MBA for administrative support for this project.

## REFERENCES

1. Carroll CL, Spinella PC, Corsi JM, Stoltz P, Zucker AR. Emergent endotracheal intubations in children: be careful if it's late when you intubate. *Pediatr Crit Care Med* 2010;11(3):343-348.
2. Nishisaki A, Ferry S, Colborn S, Defalco C, Dominguez T, Brown CA 3rd, et al; for the National Emergency Airway Registry (NEAR) and National Emergency Airway Registry for Kids (NEAR4KIDS) Investigators. Characterization of tracheal intubation process of care and safety outcomes in a tertiary pediatric intensive care unit. *Pediatr Crit Care Med* 2012;13(1):e5-e10.
3. Nishisaki A, Nguyen J, Colborn S, Watson C, Niles D, Hales R, et al. Evaluation of multidisciplinary simulation training on clinical performance and team behavior during tracheal intubation procedures in a pediatric intensive care unit. *Pediatr Crit Care Med* 2011; 12(4):406-414.
4. Nishisaki A, Donoghue AJ, Colborn S, Watson C, Meyer A, Brown CA 3rd, et al. Effect of just-in-time simulation training on tracheal intubation procedure safety in the pediatric intensive care unit. *Anesthesiology* 2010;113(1):214-23.
5. Easley RB, Segeleon JE, Haun SE, Tobias JD. Prospective study of airway management of children requiring endotracheal intubation before admission to a pediatric intensive care unit. *Crit Care Med* 2000;28(6):2058-2063.
6. Sagarin MJ, Chiang V, Sakles JC, Barton ED, Wolfe RE, Vissers RJ, Walls RM; National Emergency Airway Registry (NEAR) investigators. Rapid sequence intubation for pediatric emergency airway management. *Pediatr Emerg Care* 2002;18(6):417-423.
7. Chen JJ, Susetio L, Chao CC. Oral complications associated with endotracheal general anaesthesia. *Ma Zui Xue Za Zhi* 1990;28(2): 163-169.
8. Gomes Cordeiro AM, Fernandes JC, Troster EJ. Possible risk factors associated with moderate or severe airway injuries in children who underwent endotracheal intubation. *Pediatr Crit Care Med* 2004; 5(4):364-368.
9. Skapik JL, Pronovost PJ, Miller MR, Thompson DA, Wu AW. Pediatric safety incidents from an intensive care reporting system. *J Patient Saf* 2009;5(2):95-101.
10. Barsuk JH, McGaghie WC, Cohen ER, O'Leary KJ, Wayne DB. Simulation-based mastery learning reduces complications during central venous catheter insertion in a medical intensive care unit. *Crit Care Med* 2009;37(10):2697-2701.
11. Knudson MM, Khaw L, Bullard MK, Dicker R, Cohen MJ, Staudenmayer K, et al. Trauma training in simulation: translating skills from SIM time to real time. *J Trauma* 2008;64(2):255-263.
12. Wayne DB, Didwania A, Feinglass J, Fudala MJ, Barsuk JH, McGaghie WC. Simulation-based education improves quality of care

- during cardiac arrest team responses at an academic teaching hospital: a case-control study. *Chest* 2008;133(1):56-61.
13. Moorthy K, Munz Y, Forrest D, Pandey V, Undre S, Vincent C, Darzi A. Surgical crisis management skills training and assessment: a simulation[corrected]-based approach to enhancing operating room performance. *Ann Surg* 2006;244(1):139-147.
  14. Yee B, Naik VN, Joo HS, Savoldelli GL, Chung DY, Houston PL, et al. Nontechnical skills in anesthesia crisis management with repeated exposure to simulation-based education. *Anesthesiology* 2005;103(2):241-248.
  15. Esmail R, Cummings C, Dersch D, Duchscherer G, Glowa J, Liggett G, Hulme T; Patient Safety and Adverse Events Team. Using health-care failure mode and effect analysis tool to review the process of ordering and administering potassium chloride and potassium phosphate. *Healthc Q* 2005;8(Spec No):73-80.
  16. Duwe B, Fuchs BD, Hansen-Flaschen J. Failure mode and effects analysis application to critical care medicine. *Crit Care Clin* 2005;21(1):21-30.
  17. Marx DA, Slonim AD. Assessing patient safety risk before the injury occurs: an introduction to sociotechnical probabilistic risk modelling in health care. *Qual Saf Health Care* 2003;12(Suppl 2):ii33-ii38.
  18. Fletcher G, Flin R, McGeorge P, Glavin R, Maran N, Patey R. Anaesthetists' Non-Technical Skills (ANTS): evaluation of a behavioural marker system. *Br J Anaesth* 2003;90(5):580-588.
  19. Holcomb JB, Dumire RD, Crommett JW, Stamateris CE, Fagert MA, Cleveland JA, et al. Evaluation of trauma team performance using an advanced human patient simulator for resuscitation training. *J Trauma* 2002;52(6):1078-1085.
  20. Lockyer J, Singhal N, Fidler H, Weiner G, Aziz K, Curran V. The development and testing of a performance checklist to assess neonatal resuscitation megacode skill. *Pediatrics* 2006;118(6):e1739-e1744.
  21. Donoghue AJ, Durbin DR, Nadel FM, Stryjewski GR, Kost SI, Nadkarni VM. Effect of high-fidelity simulation on Pediatric Advanced Life Support training in pediatric house staff: a randomized trial. *Pediatr Emerg Care* 2009;25(3):139-144.
  22. Donoghue A, Ventre K, Boulet J, Brett-Fleegler M, Nishisaki A, Overly F, Cheng A; for the EXPRESS Pediatric Simulation Research Investigators. Design, implementation, and psychometric analysis of a scoring instrument for simulated pediatric resuscitation: a report from the EXPRESS pediatric investigators. *Simul Healthc* 2011;6(2):71-77.
  23. Boulet JR, Murray DJ. Simulation-based assessment in anesthesiology: requirements for practical implementation. *Anesthesiology* 2010;112(4):1041-1052.
  24. O'Donnell CP, Kamlin CO, Davis PG, Morley CJ. Endotracheal intubation attempts during neonatal resuscitation: success rates, duration and adverse effects. *Pediatrics* 2006;117(1):e16-e21.
  25. Accreditation Council for Graduate Medical Education. ACGME program requirements for graduate medical education in pediatrics. Effective July 2007. [http://www.acgme.org/acWebsite/downloads/RRC\\_progReq/320\\_pediatrics\\_07012007.pdf](http://www.acgme.org/acWebsite/downloads/RRC_progReq/320_pediatrics_07012007.pdf). Accessed April 25, 2012.
  26. Nishisaki A, Nadkarni VM, Berg RA. Pediatric advanced airway management training for non-anesthesia residents. In: Vincent JL, editor. *Yearbook of intensive care and emergency medicine*. Springer-Verlag; 2009:322-331.
  27. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. *Med Educ* 2010;44(1):50-63.
  28. Nishisaki A, Keren R, Nadkarni V. Does simulation improve patient safety? Self-efficacy, competence, operational performance, and patient safety. *Anesthesiol Clin* 2007;25(2):225-236.