Pediatric Airway Assessment Tool (PAAT): A Rating Tool to Assess Resident Proficiency in Simulated Pediatric Airway Skills Performance

Robyn Wing, MD, MPH*, Janette Baird, PhD, Susan Duffy, MD, MPH, Linda Brown, MD, MSCE, Frank Overly, MD, Mariann Nocera Kelley, MD, Chris Merritt, MD, MHPE

*Corresponding Author: rwing901@gmail.com

Abstract

Introduction: The Accreditation Council for Graduate Medical Education has identified the need for assessment of core skills for pediatric and emergency medicine residents, which includes pediatric airway management. Although there are standard courses for pediatric airway management, there is no validated tool to assess basic and advanced pediatric airway skills performance. Our objective was to develop a simulation-based tool for the formative assessment of resident pediatric airway skills performance that was concise, yet comprehensive, and to evaluate the evidence supporting the argument for the tool's validity. Methods: We developed a pediatric airway assessment tool (PAAT) to assess six major domains of pediatric airway skills performance: basic airway maneuvers, airway adjuncts, bag-valve mask ventilation, advanced airway equipment preparation, direct laryngoscopy, and video laryngoscopy. This tool consisted of a 72-item pediatric airway skills assessment checklist to be used in simulation. We enrolled 12 subjects at four different training levels to participate. Assessment scores were rated by two independent expert raters. Results: The interrater agreement was high, ranging from 0.92 (adult bagging rate) to 1 (basic airway maneuvers). There was a significant trend of increasing scores with increased training level. Discussion: The PAAT demonstrated excellent interrater reliability and provided evidence of the construct's validity. Although further validation of this assessment tool is needed, these results suggest that the PAAT may eventually be useful for assessment of resident proficiency in pediatric airway skills performance.

Keywords
Pediatric Airway, Simulation, Physician, Emergency Medicine, Pediatric Emergency Medicine, Pediatrics, Clinical Skills Assessment/OSCEs, Clinical/Procedural Skills Training, Assessment

Educational Objectives

By the using this assessment, facilitators will be able to evaluate learners:

1. Demonstrating basic airway skills maneuvers.
2. Describing and properly using basic airway equipment such as the oropharyngeal airway, nasopharyngeal airway, and bag valve mask (BVM).
3. Demonstrating effective BVM ventilation on infant and adult patients.
4. Preparing for endotracheal intubation.
5. Demonstrating direct laryngoscopy and video laryngoscopy (Glidescope) for endotracheal intubation of infant and adolescent/adult patients.

Introduction

In teaching hospitals, residents are often called to respond to children with respiratory failure, and must be proficient in basic airway management, and may be called upon to perform more advanced airway skills with supervision. However, residents are exposed to relatively few critically ill children, so must rely on alternative methods of developing proficiency in airway skills performance. Pediatric advanced life support (PALS) courses and other traditional learning methods have variable effects on proficiency. Knowledge retention and skill performance, including airway management, are reported to deteriorate quickly if not used. Medical simulation can be utilized to teach pediatric airway skills and to assess residents' abilities. The Accreditation Council for Graduate Medical Education (ACGME)
has mandated the assessment of specific core skills, including emergency management and critical care procedures,10 and advocates for the use of simulation for both instruction and assessment.11,12 Recommendations that serve as a guide for the implementation of simulation-based assessments across the domains of validity, reproducibility, feasibility, educational and catalytic effects, have recently been published.13

There is currently no complete formative assessment tool for basic and advanced pediatric airway skills performance with sufficient evidence of validity. There have been other published tools for assessment of pediatric airway skills, but these were missing important components such as airway adjuncts and video laryngoscopy.14-17 A recent publication referenced a comprehensive airway assessment checklist for adult patients, but did not include a pediatric assessment and included no validity evidence.18 Given the significant anatomic differences but did not include a pediatric assessment and included no comprehensive airway assessment checklist for adult patients, this adult assessment was not suitable to assess providers skills in pediatric patients. Our objective was to develop a comprehensive simulation-based tool for the formative assessment of resident pediatric airway skills performance and to evaluate the evidence for the tool’s validity. The construct being measured was an individual resident’s ability to manage the airway of a patient in respiratory distress or failure. This tool’s intended use was to assess whether residents possess the minimum basic skills to perform critical airway management under direct expert supervision during actual clinical encounters. Our goal was to use this tool as a formative assessment to provide feedback on current learner performance and help the learner to improve prior to performing these skills in a clinical setting. This formative assessment was one step in the process of preparing learners for eventual independent performance of the task.

We present our tool, the pediatric airway assessment tool (PAAT), and the results of a pilot study. This tool will advance assessment practice by allowing for standardized assessment of these critical skills in a safe learning environment prior to allowing learners to perform these activities in clinical practice.

**Methods**

**Tool Development**

Our team of pediatric emergency medicine specialists developed the PAAT after extensive literature review using the PALS curriculum19 and standard pediatric airway management knowledge. This tool consists of a 72-item observational checklist (Appendix A), including assessment of six domains of pediatric airway skills performance: basic airway maneuvers (three items), airway adjuncts (seven items), bag valve mask (BVM) ventilation (eight items), advanced airway equipment preparation (10 items), direct laryngoscopy (22 items), and video laryngoscopy (22 items). These six domains were categorized into three sections (ventilatory management skills, infant advanced airway, and adolescent advanced airway) on the checklist for feasibility of testing using specific airway equipment, task trainers, and manikins. Checklist items were chosen to assess critical actions in airway skills performance. The PAAT differs from previously published airway assessments in that it is concise yet comprehensive, including assessments of BVM, airway adjuncts, direct laryngoscopy, and video laryngoscopy on both infant- and adolescent-sized manikins.

A panel of 10 experts in emergency medicine, critical care, and medical education reviewed the tool for content, and to maximize the likelihood that performance of the critical actions measured in the test setting could be extrapolated to performance in clinical practice. Reviewers confirmed that questions and checklist items were accurate and expected for competent pediatric airway skills performance, eliminated inappropriate items, and added any necessary missing items. Modifications to the response process were made based on this feedback. Checklist items were carefully worded to ensure standardization of assessment and facilitate scoring, with scripted prompts to be used by the rater whenever possible. Two physician raters (who were part of the tool development team) trained for familiarity with the checklist and were directed to follow the assessment tool explicitly when rating the performance of the subjects enrolled in the pilot study.

The equipment needed included:

- **Manikins:**
  - SimMan (or other adolescent/adult-sized manikin that allowed for placement of airway adjuncts, endotracheal intubation, and assessment of chest rise with BVM)
  - SimBaby (or other infant-sized manikin that allowed for placement of airway adjuncts, endotracheal intubation, and assessment of chest rise with BVM)

- **Equipment:**
  - Stethoscope
  - Towels
  - BVMs with masks—multiple sizes*
  - Nasopharyngeal airways—all sizes*
  - Oropharyngeal airways—all sizes*
  - Glidescope with various size handles*
  - Glidescope stylet
  - Laryngoscope blades—multiple sizes*

---

*refer to the original document for specific sizes.*
Pilot Subject Enrollment and Sessions
We recruited a convenience sample of 12 volunteer participants of different training levels (equal numbers of medical students, residents, pediatric emergency medicine fellows, and pediatric emergency medicine attending physicians) to pilot the assessment tool. Subjects gave verbal informed consent and completed a pretest to assess their pediatric airway management experience and self-efficacy with regard to pediatric airway skills. Subjects then entered the simulation lab where two high-fidelity simulation manikins (Laerdal SimBaby and SimMan, Laerdal Corporation) were set up on stretchers. All airway equipment, as listed above, was displayed for use and the participant was expected to choose the proper size equipment for the manikin, as part of demonstrating their skill level in pediatric airway management. Participants performed the tasks listed in the pediatric airway skills assessment checklist (Appendix A) individually with scripted prompting from our raters and were instructed to verbalize their actions. Raters asked the participants to perform tasks using the outlined prompts in the PAAT, such as, “Demonstrate how to properly position and open the airway of this patient presenting in respiratory distress.” Raters were instructed to immediately remediate the participant prior to proceeding to the next section if they did not successfully complete an item. This remediation was initially done with additional scripted prompting, such as, “Is there anything else that you can do with your hands to help open the airway of a patient?” and then through directed instruction to the participant if the scripted prompt did not lead to corrective action.

Scoring
Numerical scoring was performed solely for the purposes of validity testing of this formative tool. The two raters independently rated the individual performance of all 12 subjects in real time using the observational PAAT checklist, rating items as 0 if subjects did not successfully complete the task without additional prompting, or 1 if subjects completed the task successfully on their first attempt. For the purpose of this pilot study and validity testing, only participants’ first attempts were scored. Final scores were the average of the two raters’ scores across items. Three summary scores were reported. The total basic score included scores for basic airway maneuvers, airway adjunct use, and bag-valve mask ventilation. The total advanced score included advanced airway equipment preparation, direct laryngoscopy, and video laryngoscopy. The overall skill score was the sum of basic and advanced scores.

For use of this formative assessment tool in practice, it was decided that in order to successfully finish the activity, the participant would have to successfully complete the entire checklist. Participants would be required to correct their errors if they inserted any adjunct in a manner that would be dangerous to a patient. Our group devoted considerable time to the issue of appropriate context for use of PAAT, which ultimately was decided by consensus of our expert panel. Because our intended use was to measure the basic pediatric airway skills such that a learner would be able to perform under direct supervision, we determined that the standard should allow for coached performance. We trained our raters to provide real-time prompts, followed by immediate feedback and one-step remediation if a learner was unsuccessful at any step. However, we did agree as a group that because airway management is such a critical skill and that failure carries life-threatening consequences, each of the steps on the PAAT checklist was critical to safe and effective performance and that inability to achieve any of the steps after real-time remediation would require substantial reteaching of the skills. In this sense, then, failure to perform any step of the process could be deemed failure of the process. Trainees who were unable to complete the task even with coaching would be offered additional training and support on an individual basis.

Statistical Analyses
Each rater scored each checklist component as present (1) or absent (0). The scores from each rater across participants were summed, and the average scores for basic, advanced, and overall airway skills between the two raters were used in the analyses to assess differences in scores as a function of experience level across the four groups (medical students, residents, pediatric emergency medicine fellows, and pediatric emergency medicine attending physicians). Scores were used to calculate agreement rates and interrater reliability. Data were analyzed using SAS Software (version 9.4). Comparisons between the groups were evaluated using the Kruskal-Wallace, and the Jonckheere-Terpstra test assessed the trend in increasing differences between the experience groups.

The agreement between raters for the three summary scores was calculated in three ways. First, a Pearson correlation coefficient for each overall skill score and the composite scores for the basic and advanced totals was calculated. Second, a Cohen’s Kappa (K) coefficient was calculated as a measure of interrater reliability.
Finally, a Kappa squared ($K^2$), was used to provide an estimate of the accuracy (by the percentage of variance estimated by the agreement between raters) of these scores.\textsuperscript{20} Cohen's Kappa and $K^2$ were calculated only for the summary scores and not the basic or advanced scores, as the restricted variance within the composite elements would likely inflate the Kappa values. The 95% confidence interval on rater agreement was also reported.

**Results**

**Participant Characteristics**

There was a broad range of prior pediatric airway experience ranging from students (no prior experience) to experts. Pediatric emergency medicine fellows and attending physicians had more experience in most domains, notably in that of video laryngoscopy, where no medical students or residents had any prior experience in this skill.

**Validity Evidence**

The PAAT was designed to provide a comprehensive formative assessment for resident pediatric airway skills performance. The process of collecting and interpreting validity evidence to support reliance on an assessment's decision was fundamental to its use.\textsuperscript{21} Sources of validity evidence were summarized in Table 1.\textsuperscript{22} Kane's framework identifies four inferences used to support the validity argument: scoring, generalization, extrapolation, and implications.\textsuperscript{23-25} The validity argument rested on the intended use of the assessment tool. Our tool was intended to measure and provide feedback to learners such that a resident would be sufficiently prepared to perform these skills under supervision in clinical practice.

The PAAT was intended to generate reproducible, accurate scoring through use of scripted instructions, clear checklist items, and a simple scoring rubric. Generalization—using a score as a reflection of actual performance in a test setting—was demonstrated with strong interrater reliability.\textsuperscript{26} The correlation between raters’ evaluation of the elements within each summary score were high, ranging from 0.92 (adolescent bagging rate) to 1 (basic airway maneuvers) (Table 2). Interrater reliability ($K$) across the summary scores total was high (0.92 to 1). The $K^2$ for each summary score (1 total basic score; 0.83 total advanced score; 0.85 total score) indicated little to no variance in the observed airway skills performance due to disagreement between the raters.

Extrapolation of scores to be used as a reflection of real-world performance was supported through demonstrating content evidence provided by expert panel review and derivation from

### Table 1. Sources of Validity Evidence for the Proposed Pediatric Airway Assessment Tool

| Area of Validity                     | Evidence                                                                 |
|--------------------------------------|--------------------------------------------------------------------------|
| Content                              | Content of knowledge relates to intended standards/objectives.Tool developers are themselves content experts.Additional content experts' review.Simulated cases have undergone case development by content experts. |
| Response process                     | Simple scoring rubric and scripted checklist minimizes error.Accuracy and reproducibility of assessment across pilot participants.Scores are easily understood by raters and trainees alike.Scripted prompts provided by assessors vetted by content experts to assure that all residents undergo identical assessment.Common simulation standards adhered to, so as to prevent variability. |
| Internal structure                   | Score reporting reliability metrics.High inter-rater reliability of checklist component.Low rater error observations.High discriminant validity. |
| Relationship to other variables      | Simulation in other airway assessment settings has relatively high generalizability.Construct validity demonstrated by a significant trend toward increasing scores with increased training level. |
| Consequence                          | Low-stakes, formative assessment. |

### Table 2. Scoring Comparison Between Raters

| No. of Items | Correlation Coefficient (95% CI) | Cohen's Kappa (95% CI) |
|--------------|----------------------------------|------------------------|
| Total score  | 0.99 (0.96, 1.00)                | 1.00 (1.00, 1.00)      |
| Basic maneuvers | 1.00 (0.99, 1.00)            |                        |
| Airway adjuncts | 0.99 (0.96, 1.00)            |                        |
| Bag-valve mask ventilation | 0.95 (0.83, 0.99)        |                        |
| Adult bagging rate | 0.93 (0.76, 0.98)         |                        |
| Total advanced score | 0.99 (0.99, 1.00) | 0.91 (0.88, 0.94) |
| Advanced airway equipment | 0.94 (0.79, 0.98) |                        |
| Direct laryngoscopy | 0.99 (0.98, 1.00) |                        |
| Video laryngoscopy | 1.00 (0.99, 1.00) |                        |
| Overall skill score | 1.00 (0.99, 1.00) | 0.92 (0.89, 0.95) |

Abbreviation: CI, confidence interval
\textsuperscript{*}Not calculated into total basic score.
As seen in Table 3, there was a significant effect of experience level on the knowledge component and in each pediatric airway score (p = .03). For the procedural scores, there was also an ordering effect, with higher scores seen with increased levels of training (p = .003), with most differences noted between fellows/attendings (i.e., experienced) and residents/medical students (i.e., inexperienced). Fellows and attending physicians outscored medical students and residents in all domains. Fellows and attending physicians had similar scores for knowledge and performance on basic airway skills and overall score, and fellows had the highest total advanced airway score. 

### Discussion

Formative assessment-based frameworks focus on training for physician performance and patient care outcomes, emphasizing assessment and correction of observed, discrete, measurable behaviors.27-29 The ACGME has established benchmarks, which include the assessment of proficiency in airway management.10 However, developing rigorous assessment tools remains challenging. Tools should be reliable and predictive of future success, with clearly defined outcomes so that observers can make objective judgments.30 Tools have been published to assess procedural skills including central venous catheter insertion,31 pediatric resuscitation,14 and lumbar puncture.32,33 However, validity evidence for tools to assess pediatric airway skills performance is lacking. Therefore, we designed the PAAT as a formative assessment to prepare learners in their progress towards competency in these skills. We, refined it based on expert feedback and prior airway checklists, and then piloted it to collect and interpret validity evidence for its future use.15-17

The tool may be modified into two separate tools—one for basic airway assessment and one for advanced airway assessment—depending on the intended use of the tool. For the purposes of scoring and testing the tool, we only scored the first pass, but we recommend rapid cycle repetition and allowing several attempts as an educational effort with future uses.

One interesting finding was that fellows scored higher than attendings in the advanced airway score. The reason for this difference was likely multifactorial. One possibility included the fact that the fellows had more recent training on, and experience with, advanced airway skills, particularly video laryngoscopy. These assessments strengthened the inference that PAAT scoring ought to reflect performance in real-life scenarios. We believe that the strength of the scoring, generalization and extrapolation inferences supported further investigation of the tool's use in assessment of pediatric airway skills performance.

We presented several aspects of validity evidence for our tool which was piloted and is expected to be used for training in the simulation environment. With any assessment tool, there are likely to be threats—anticipated or unanticipated—to the validity inferences between observed performance in the test setting and the unobserved proficiency in the setting of actual practice. While we sought to minimize the impact of these threats, their potential must be acknowledged. Not included in the current version of the PAAT is a separate knowledge assessment. It is likely that there are certain discrete knowledge items (e.g., appropriate depth of tube insertion) that could be assessed using a multiple-choice questionnaire or other format.

There were a number of limitations to our tool and this study. The relatively small size of the pilot sample was a clear limitation. While we detected significant differences among training groups,

---

### Table 3. Pediatric Airway Assessment Scores (n = 12, 95% CI)

| Category                   | Medical Student (M, 95% CI) | Resident (M, 95% CI) | PEM Fellow (M, 95% CI) | PEM Attending (M, 95% CI) | P    | P    |
|---------------------------|-----------------------------|---------------------|------------------------|--------------------------|------|------|
| Total basic score         | 8.7 (0, 18.7)               | 14.5 (9.1, 19.9)    | 16.7 (14.7, 18.6)      | 16 (15.2, 18.1)          | .03  | .003 |
| Basic airway maneuvers    | 1.3 (0, 2.8)                | 2 (0, 4.5)          | 4 (0, 8.3)             | 3.7 (0, 8.3)             |      |      |
| Airway adjuncts           | 2.3 (0, 8.5)                | 6 (3.5, 8.5)        | 6.2 (4.3, 8.1)         | 6.3 (4.9, 7.8)           |      |      |
| Bag-valve mask ventilation| 5 (2.5, 7.5)                | 6.5 (5.3, 7.7)      | 7.5 (6.3, 8.7)         | 7.7 (6.2, 9.1)           |      |      |
| Total advanced score      | 10 (0, 34.5)                | 23.2 (21.7, 25.1)   | 52.3 (49.7, 55.2)      | 47.8 (35.3, 60.4)        | .02  | .001 |
| Advanced airway equipment | 0.8 (0, 3.4)                | 5 (2.9, 7.2)        | 9 (5.7, 12.3)          | 7.2 (4.6, 9.8)           |      |      |
| Direct laryngoscopy       | 9 (0, 30.2)                 | 18.7 (14.9, 22.5)   | 21.7 (21, 22.4)        | 20.5 (14.1, 27)          |      |      |
| Video laryngoscopy        | 0 (0)                       | 0 (0)               | 21.7 (20.23, 23.10)    | 20.2 (16.58, 23.75)      |      |      |
| Overall skill score       | 18.7 (0, 51.8)              | 37.7 (31.4, 43.9)   | 69 (65.3, 72.7)        | 64.5 (50.5, 78.5)        | .02  | .001 |

**Abbreviation:** CI, confidence interval; PEM, pediatric emergency medicine

*Significance value for Kruskal-Wallace test.

aSignificance value for Jonckheere-Terpstra test.
larger studies should determine more clearly how the tool performs, particularly for learners away from the extremes of experience. Additionally, the volunteerism of the pilot sample may select for more conscientious participants, which may actually affect skill set. The use of raters who were part of the development team, not blinded to the subjects' training level, could introduce bias and limit the interpretation of inter-rater reliability. However, this will be likely during routine use of this tool at any institution. To standardize the assessment and minimize bias, each scenario was tightly scripted to ensure consistency. Checklist items were observable, objective, yes/no questions. Still, the tool may not perform as well if used by educators who have not been as extensively trained. Due to the use of a convenience sample, selection bias may have occurred, though participants were selected based only on availability, which should have no bearing on pediatric airway skill set. The use of simulation versus live patients for assessment is necessary but does limit extrapolation to live clinical performance. If the physical, emotional, or conceptual realism of the scenario were not maintained, this may affect the response process and reliability of the tool. We cannot comment on the ability of this assessment to predict performance in the clinical setting based on performance in the simulated setting. This should be addressed in future studies.

Conclusion
Formative assessment tools are a crucial component to address the mandate to assess patient care procedural skills during residency training. Such tools provide objective measurements of resident procedural skills that can be used by programs to enhance training by tracking resident progress over time and guiding specific changes in program curricula. Prior to the development of this resource, there existed no validated formative assessment tool for training in comprehensive pediatric airway skills performance. This publication provides both an easily implemented assessment tool for training and initial evidence to support validity of its use.

Appendices
A. Pediatric Airway Skills Assessment Checklist.docx

All appendices are peer reviewed as integral parts of the Original Publication.

Robyn Wing, MD, MPH: Assistant Professor, Departments of Emergency Medicine & Pediatrics, Division of Pediatric Emergency Medicine, Alpert Medical School of Brown University and Rhode Island Hospital/Hasbro Children's Hospital; Director of Pediatric Simulation, Lifespan Medical Simulation Center
Janette Baird, PhD: Associate Professor, Department of Emergency Medicine and Injury Prevention Center, Alpert Medical School of Brown University
Susan Duffy, MD, MPH: Professor, Departments of Emergency Medicine & Pediatrics, Division of Pediatric Emergency Medicine, Alpert Medical School of Brown University and Rhode Island Hospital/Hasbro Children's Hospital
Linda Brown, MD, MSCE: Associate Professor, Departments of Emergency Medicine & Pediatrics, Division of Pediatric Emergency Medicine, Alpert Medical School of Brown University and Rhode Island Hospital/Hasbro Children's Hospital; Vice Chair of Pediatric Emergency Medicine; Director of the Lifespan Medical Simulation Center
Frank Overly, MD: Professor, Departments of Emergency Medicine & Pediatrics, Division of Pediatric Emergency Medicine, Alpert Medical School of Brown University and Rhode Island Hospital/Hasbro Children's Hospital; Medical Director of Hasbro Emergency Department
Mariann Nocera Kelley, MD: Assistant Professor, Departments of Pediatrics and Emergency Medicine/Traumatology, Division of Pediatric Emergency Medicine, University of Connecticut School of Medicine, Connecticut Children’s Medical Center; Director of Simulation Education, University of Connecticut School of Medicine
Chris Merritt, MD, MHPE: Associate Professor, Departments of Emergency Medicine & Pediatrics, Division of Pediatric Emergency Medicine, Alpert Medical School of Brown University and Rhode Island Hospital/Hasbro Children's Hospital; Director, Brown Emergency Medicine Medical Education Research Fellowship

Acknowledgments
We gratefully acknowledge the additional members of our expert panel: Lee Polikoff, MD, Brian Clyne, MD, David Lindquist, MD, and Leo Kobayashi, MD.

Disclosures
None to report.

Funding/Support
None to report.

Ethical Approval
This project was approved by the Rhode Island Hospital Institutional Review Board, with updated approval on February 2, 2018, Project #833391-9.

References
1. Chen EH, Cho CS, Shofe FS, Mills AM, Baren JM. Resident exposure to critical patients in a pediatric emergency department. Pediatr Emerg Care. 2007;23(11):774-778. https://doi.org/10.1097/pec.0b013e318159ffef

Copyright © 2020 Wing et al. This is an open-access publication distributed under the terms of the Creative Commons Attribution license.
2. Kendirli T, Caltik A, Duman M, et al. Effect of pediatric advanced life support course on pediatric residents’ intubation success. *Pediatr Int*. 2011;53(1):94-99. https://doi.org/10.1111/j.1442-200x.2010.03128.x

3. Stringer KR, Bajenov S, Yentis SM. Training in airway management. *Anaesthesia*. 2002;57(10):967-983. https://doi.org/10.1046/j.1365-2044.2002.02830.x

4. Nadel FM, Lavelle JM, Fein JA, Giardino AP, Decker JM, Durbin DR. Teaching resuscitation to pediatric residents: the effects of an intervention. *Arch Pediatr Adolesc Med*. 2000;154(10):1049-1054. https://doi.org/10.1001/archpedi.154.10.1049

5. Mancini ME, Kaye W. The effect of time since training on house officers’ retention of cardiopulmonary resuscitation skills. *Am J Emerg Med*. 1985;3(1):31-32. https://doi.org/10.1016/0735-6757(85)90008-7

6. Gass DA, Curry L. Physicians’ and nurses’ retention of knowledge and skill after training in cardiopulmonary resuscitation. *Can Med Assoc J*. 1983;128(5):550-551.

7. Overly FL, Sudikoff SN, Shapiro MJ. High-fidelity medical simulation as an assessment tool for pediatric residents’ airway management skills. *Pediatr Emerg Care*. 2007;23(1):11-15. https://doi.org/10.1097/pec.0b013e31802c61d5

8. Sudikoff SN, Overly FL, Shapiro MJ. High-fidelity medical simulation as a technique to improve pediatric residents’ emergency airway management and teamwork: a pilot study. *Pediatr Emerg Care*. 2009;25(10):651-656. https://doi.org/10.1097/pec.0b013e3181bd93ea

9. Kennedy CC, Cannon EK, Warner DO, Cook DA. Advanced airway management simulation training in medical education: a systematic review and meta-analysis. *Crit Care Medicine*. 2014;42(1):169-178. https://doi.org/10.1097/ccm.0b013e31829a721f

10. ACGME Program Requirements for Graduate Medical Education in Pediatrics. Accreditation Council for Graduate Medical Education. *Published 2020*. https://www.acgme.org/Portals/0/PFAAssets/ProgramRequirements/320_pediatrics_2020.pdf

11. Hamstra S, Philibert I. Simulation in graduate medical education: understanding uses and maximizing benefits. *J Grad Med Educ*. 2012;4(4):539-540. https://doi.org/10.4300/JGME-D-12-00260.1

12. Mills DM, Williams DC, Dobson JV. Simulation training as a mechanism for procedural and resuscitation education for pediatric residents: a systematic review. *Hosp Pediatr*. 2013;3(2):167-176. https://doi.org/10.1542/hped.2012-0041

13. Hall AK, Chaplin T, McColl T, et al. Harnessing the power of simulation for assessment: consensus recommendations for the use of simulation-based assessment in emergency medicine. *CJEM*. 2020;22(2):193-203. https://doi.org/10.1017/cem.2019.488

14. Brett-Fleegler MB, Vinci RJ, Weiner DL, Harris SK, Shih MC, Kleinman ME. A simulator-based tool that assesses pediatric resident resuscitation competency. *Pediatrics*. 2008;121(3):e597-e603. https://doi.org/10.1542/peds.2005-1259

15. Faudeux C, Tran A, Dupont A, et al. Development of reliable and validated tools to evaluate technical resuscitation skills in a pediatric simulation setting: resuscitation and emergency simulation checklist for assessment in pediatrics. *J Pediatr*. 2017;188:252-257.e6. https://doi.org/10.1016/j.jpeds.2017.03.055

16. Way DP, Panchal AR, Finnegan GI, Terndrup TE. Airway management proficiency checklist for assessing paramedic performance. *Prehosp Emerg Care*. 2017;21(3):354-361. https://doi.org/10.1080/10903127.2016.1263368

17. Sawyer T, Creamer K, Punet R, et al. Pediatric procedural skills training curriculum. *MedEdPORTAL*. 2010;6:8094. https://doi.org/10.15766/medp_2374-8265.8094

18. Galway UA, Straker T, Foley L, Aziz M, Woodworth G. Anesthesia residency training in airway management: a competency-based model curriculum. *A A Pract*. 2019;13(5):197-199. https://doi.org/10.1213/xaa.0000000000001046

19. American Heart Association. *Pediatric Advanced Life Support Provider Manual*. American Heart Association; 2011.

20. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276-282.

21. Andreatta PB, Gruppen LD. Conceptualising and classifying validity evidence for simulation. *Med Educ*. 2009;43(11):1028-1035. https://doi.org/10.1111/j.1365-2923.2009.03454.x

22. Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ*. 2003;37(9):830-837. https://doi.org/10.1046/j.1365-2923.2003.01594.x

23. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane’s framework. *Med Educ*. 2015;49(6):560-575. https://doi.org/10.1111/medu.12678

24. Kane MT. Validation. In: Brennan RL, ed. *Educational Measurement*. Praeger Publishers; 2006:17-64.

25. Kane MT. Validating the interpretations and uses of test scores. *J Educ Meas*. 2013;50(1):1-73.

26. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37(5):360-363.

27. Carraccio C, Wolfsthal SD, Englander R, Fentzer K, Martin C. Shifting paradigms: from Flexner to competencies. *Acad Med*. 2002;77(5):361-367. https://doi.org/10.1097/00001888-200205000-00003

28. Hawkins RE, Welcher CM, Holmboe ES, et al. Implementation of competency-based medical education: are we addressing the concerns and challenges? *Med Educ*. 2015;49(11):1086-1102. https://doi.org/10.1111/medu.12831

29. Frank JR, Snell LS, Cate OT, et al. Competency-based medical education: theory to practice. *Med Teach*. 2010;32(8):638-645. https://doi.org/10.3109/0142159x.2010.501190
30. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65(9 Suppl):S63-S67. https://doi.org/10.1097/00001888-199009000-00004

31. Barsuk JH, Cohen ER, Potts S, et al. Dissemination of a simulation-based mastery learning intervention reduces central line-associated bloodstream infections. *BMJ Qual Saf*. 2014; 23(9):749-756. https://doi.org/10.1136/bmjqs-2013-002665

32. Gerard JM, Kessler DO, Braun C, Mehta R, Scalzo AJ, Auerbach M. Validation of global rating scale and checklist instruments for the infant lumbar puncture procedure. *Simul Healthc*. 2013;8(3):148-154. https://doi.org/10.1097/sih.0b013e3182802d34

33. Iyer MS, Santen SA, Nypaver M, et al. Assessing the validity evidence of an objective structured assessment tool of technical skills for neonatal lumbar puncture. *Acad Emerg Med*. 2013;20(3):321-324. https://doi.org/10.1111/acem.12093

Received: July 16, 2019
Accepted: June 4, 2020
Published: October 19, 2020