Development of a Tool to Assess Basic Competency in the Performance of Rigid Bronchoscopy

Kamran Mahmood[1], Momen M. Wahidi[1], Kathryn E. Osann[2], Kathleen Coles[1], Scott L. Shofer[1], Ellen E. Volker[1], and Mohsen Davoudi[3]

[1]Department of Medicine, Division of Pulmonary, Allergy, and Critical Care Medicine, Duke University Medical Center, Durham, North Carolina; and [2]Department of Medicine, Division of Hematology–Oncology, and [3]Division of Pulmonary and Critical Care Medicine, University of California Irvine, Orange, California

Abstract

Rationale: Rigid bronchoscopy is increasingly used by pulmonologists for the management of central airway disorders. However, an assessment tool to evaluate the competency of operators in the performance of this technique has not been developed. We created the Rigid Bronchoscopy Tool for Assessment of Skills and Competence (RIGID-TASC) to serve as an objective, competency-oriented assessment tool of basic rigid bronchoscopic skills, including rigid bronchoscopic intubation and central airway navigation.

Objectives: To assess whether RIGID-TASC scores accurately distinguish the basic rigid bronchoscopy skills of novice, intermediate, and expert operators, and to determine whether RIGID-TASC has adequate interrater reliability when used by different independent testers.

Methods: At two academic medical centers in the United States, 30 physician volunteers were selected in three categories: 10 novices at rigid bronchoscopy (performed at least 50 flexible, but no rigid, bronchoscopies), 10 operators with intermediate experience (performed 5–20 rigid bronchoscopies), and 10 experts (performed ⩾100 rigid bronchoscopies). Participants included pulmonary and critical care fellows, interventional pulmonology fellows, and faculty interventional pulmonologists. Each subject then performed rigid bronchoscopic intubation and navigation on a manikin, while being scored independently by two testers, using RIGID-TASC.

Measurements and Main Results: Mean scores for three categories (novice, intermediate, and expert) were 58.10 (±4.6 [SE]), 78.15 (±3.8), and 94.40 (±1.1), respectively. There was significant difference between novice and intermediate (20.05, 95% confidence interval [CI] = 7.77–32.33, $P = 0.001$), and intermediate and expert (16.25, 95% CI = 3.97–28.53, $P = 0.008$) operators. The interrater reliability (intraclass correlation coefficient) between the two testers was high ($r = 0.95$, 95% CI = 0.90–0.98).

Conclusions: RIGID-TASC showed evidence of construct validity and interrater reliability in this setting and group of subjects. It can be used to reliably and objectively score and classify operators from novice to expert in basic rigid bronchoscopic intubation and navigation.

Keywords: bronchoscopy; pulmonary surgical procedures; competency-based education

(Received in original form September 8, 2015; accepted in final form November 21, 2015)

Author Contributions: K.M., M.M.W., S.L.S., and M.D. designed and implemented the study; K.M., K.C., E.E.V., and M.D. collected the data; K.E.O., K.M., and M.D. were responsible for data analysis; all authors contributed to the writing of the manuscript.

Correspondence and requests for reprints should be addressed to Kamran Mahmood, M.D., M.P.H., Division of Pulmonary, Allergy and Critical Care Medicine, Department of Medicine, Duke University Medical Center, DUMC 102356, Durham, NC 27710. E-mail: k.mahmood@duke.edu

Ann Am Thorac Soc Vol 13, No 4, pp 502–511, Apr 2016

DOI: 10.1513/AnnalsATS.201509-593OC
Internet address: www.atsjournals.org

Rigid bronchoscopy, having suffered a decline in popularity with the introduction of the flexible bronchoscope, has experienced a resurgence in use by pulmonologists for therapeutic indications (1). Interventional pulmonary fellowships have been introduced in the United States, and formal training in rigid bronchoscopy is considered a backbone of these programs. However, within these training programs, there is no standard approach to performing rigid bronchoscopy, and procedural numbers vary widely from program to program (2, 3). Consequently, an objective tool to permit assessment of procedural competence is urgently needed to ensure adequate training of interventional pulmonologists completing these programs.

The American College of Chest Physicians (ACCP), American Thoracic

Society, and European Respiratory Society have recommended a minimum number of 20 rigid bronchoscopy procedures for initial competency (4, 5). This number is solely based on expert opinion, as opposed to discernible scientific evidence. Recent guidelines from ACCP have acknowledged this limitation (6). Multiple studies have shown that trainees can show significant variation in the rate at which they acquire procedural skills, leading to operators with differing skill levels at these recommended "threshold" numbers, hence concluding that performance-based metrics be incorporated into the process of competency assessment (7, 8).

To address the need for competency-based assessment of rigid bronchoscopy, we developed an objective, 23-point, checklist-based instrument, the Rigid Bronchoscopy Tool for Assessment of Skills and Competence (RIGID-TASC). This instrument assesses the basic competency of an operator in performing rigid bronchoscopic intubation and central airway navigation. We hypothesized that: (*1*) RIGID-TASC can discriminate between rigid bronchoscopy operators at novice, intermediate, and expert levels based on their intubation and central airway navigation skills, thus establishing construct validity; and (*2*) RIGID-TASC has adequate interrater reliability, generating reproducible scores when used by independent examiners.

Methods

Instrument

RIGID-TASC is a 23-point checklist that gives a score of 0–100 (Appendix). It is designed to assess proper assembly of the rigid bronchoscope, correct positioning of the patient, atraumatic introduction of the bronchoscope through the upper airway, trachea, and mainstem bronchi, number of attempts, and time required to complete the task. Every point addresses an important step in the process. A score of "1" is given based on a "yes" response, or "0" based on a "no" response. There are no partial scores, to allow for reproducible scoring and avoid subjectivity.

The score for each item in RIGID-TASC was weighted to achieve a composite score, as is standard practice (9–12) and used in development of other bronchoscopy tools, such as Bronchoscopy Skills and Task Assessment Tool (BSTAT) and Endobronchial Ultrasound Skills and Task Assessment Tool (EBUS-STAT) (13, 14). The score of each item (corresponding to each step) was based on: (*1*) the importance of that step in contributing to procedural safety and effectiveness; and (*2*) the difficulty to master that step. The specific numeric values of these weights were reached by consensus among the authors and devised before the tool was tested on the participants.

Testing Protocol

The study was considered exempt from review by the Institutional Review Board at Duke University Medical Center (Protocol no. Pro00049166; Durham, NC). A total of 30 volunteers, at novice, intermediate, and expert levels based on their rigid bronchoscopy skills, were recruited at two teaching institutions (Duke University and University of California Irvine, Orange, CA). Novice operators were defined as physicians who had performed more than 50 flexible bronchoscopies, but no rigid bronchoscopies. These novices were given a 10-minute instruction on the basic techniques of rigid bronchoscopy and bronchoscopic intubation before testing. Intermediate operators were defined as operators who had performed 5–20 rigid bronchoscopies, and at least 50 flexible bronchoscopies. The expert group was defined as operators who had performed more than 100 rigid bronchoscopies. The operators were also asked to complete a survey with their demographics, number of flexible and rigid bronchoscopies, self-assessed expertise level, and other variables.

A 10-mm (outer diameter) rigid bronchoscope (Bryan Corp.,Woburn, MA) was introduced in a manikin (Nakhosteen Bronchoscopy Model; CLA, Coburg, Germany). Two examiners independently assessed and recorded the performance of the operators using the RIGD-TASC while they were performing the procedures. Both examiners were blinded to each other's scores.

Statistical Analysis

Sample size was calculated *a priori* using nQuery 7.0 (Statistical Solutions, Boston, MA). To achieve a power of 80% and detect an effect size of 0.36 with a significance level of 0.05, 10 subjects were required at each level of expertise.

Subjects were divided into three groups, as described previously here, based on their expertise in rigid bronchoscopy: novice, intermediate, and expert. Descriptive data are shown as mean (±SD). The RIGID-TASC scores are shown as mean (±SE). Continuous variables were compared using Student's two-tailed t test, whereas categorical variables were compared with chi-square test. A P value of 0.05 or less was considered statistically significant.

The average total score between the three groups was compared with ANOVA using Tukey's method of adjustment for multiple comparisons. Interrater reliability was assessed with intraclass correlation coefficient. Agreement between self-assessed expertise and RIGID-TASC score was checked using Cohen's κ. SYSTAT version 13 (Systat Software, Inc., San Jose, CA) software was used to perform the statistical analysis.

Results

A total of 10 pulmonary and critical care fellows with no prior rigid bronchoscopy experience were recruited at the novice level. Of the 10 participants with an intermediate level of skill, 6 were interventional pulmonary fellows, whereas 4 were pulmonary and critical care fellows who had performed over 50 flexible and 5–20 rigid bronchoscopies on patients or a manikin. The mean number of rigid bronchoscopies performed by this group was 12.5, and the median was 12. The 10 experts, who had to have experience with more than 100 procedures, all stated a history of several hundred rigid bronchoscopies, but could not provide an accurate number.

The demographics of the participants are shown in Table 1. The mean score of operators in the novice, intermediate, and expert groups were 58.10 (±4.6 [SE]), 78.15 (±3.8), and 94.40 (±1.1), respectively, as shown in Figure 1. There was a significant difference between novice and intermediate (20.05, 95% confidence interval [CI] = 7.77–32.33, $P = 0.001$), and intermediate and expert (16.25, 95% CI = 3.97–28.53, $P = 0.008$) operators. The interrater reliability (intraclass correlation coefficient) between the two testers was an r value of 0.95 (95% CI = 0.90–0.98).

RIGID-TASC showed good discriminatory properties at different stages of the procedure, such as preintubation

Table 1. Demographics of operators

Category	Novice	Intermediate	Expert	*P* Value for Difference
Mean ± SD age, yr	30 ± 1	33.2 ± 3	40.7 ± 4	<0.001
Female sex, n (%)	1 (10)	3 (30)	2 (20)	0.5
Years from medical school	4.1 ± 0.3	7 ± 3	13.4 ± 5	<0.001
Current position	Pulmonary/critical care fellows	Interventional pulmonology fellows (n = 6) Pulmonary/critical care fellows (n = 4)	Faculty interventional pulmonologists	
Self-assessed skill level (%)	No experience (100)	Beginner (100)	Expert (88)	

assembly of the bronchoscope and optimal positioning of the patient, upper airway navigation, lower airway navigation in trachea and mainstem bronchi, number of attempts required, and time required to complete the procedure, as shown in Table 2 and Figure 2. There was also a good agreement between the self-assessed level of expertise and actual expertise based on the RIGID-TASC scores (Cohen's κ = 0.75, 95% CI = 0.55–0.95).

Discussion

RIGID-TASC demonstrated validity and reliability as an assessment tool of rigid bronchoscopic intubation and airway navigation, in the settings and subjects of this study. Validity was demonstrated by its ability to discriminate operators based on their skills into novice, intermediate, and expert categories. Reliability was shown by its high interrater agreement, with independent examiners generating reproducible and similar scores for operators at the same expertise level.

Rigid bronchoscopy is frequently used for management of malignant and benign central airway disorders. The advantages of using a rigid bronchoscope include the maintenance of a secure airway with effective ventilation and oxygenation while multiple rigid and flexible instruments are introduced through the bronchoscope (15, 16). Thus far, rigid bronchoscopy training has lacked an objective tool for assessment of competency and task-specific feedback. Indeed, a tool has been introduced recently for otolaryngology trainees (17, 18), yet the clinical training and scope of practice of an interventional pulmonologist are very different from those of an otolaryngologist. Hence, we designed RIGID-TASC for evaluating rigid bronchoscopy intubation and central airway navigation skills, with the goal of its being incorporated in the interventional pulmonology curriculum as an objective scoring and feedback tool.

Measuring and assessing competency is a complex task. Assessment of clinical skills can be performed using a standardized and validated checklist-based assessment tool. A checklist is a series of desired and effective behaviors or steps, which, when undertaken in a certain order, add up to a complex competency (19). Checklists are of proven educational and safety-enhancing value in a variety of fields, including complex physician–patient interactions, such as operative procedures (20–23). The BSTAT and the Endobronchial Ultrasound Skills and Task Assessment Tool are examples of checklist-based tools that have been introduced to assess the competency of the operators in flexible bronchoscopy and endobronchial ultrasound, respectively (13, 14).

An important aspect of developing a checklist-based assessment tool is to ensure that it captures the important components or steps of that complex task, and that these constituent elements are weighted appropriately (24). According to Nunnally (20), the plan and procedures of constructing measures is a primary determinant of construct validity. Therefore, the most important step in adopting a task checklist is to identify the key constituent steps in the performance of that task. We designed RIGID-TASC to precisely assess the most important steps in rigid bronchoscopic intubation and navigation, focusing our attention specifically on the safety and avoidance of trauma to the upper and lower airways (our emphasis on weighting each step primarily being safety, effectiveness, and difficulty to master that step). It retains its discriminatory qualities when the rigid bronchoscopy performance is broken down into its subcategories, such as

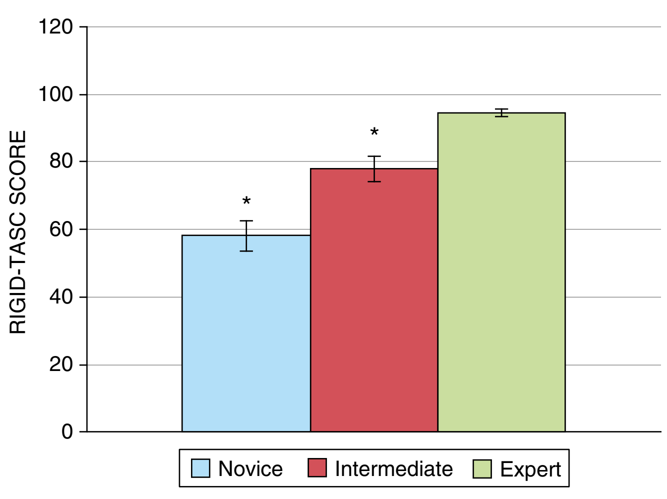

Figure 1. Rigid Bronchoscopy Tool for Assessment of Skills and Competence (RIGID-TASC) scores for novice, intermediate, and expert groups. Higher score depicts better performance. *Statistically significant difference. The RIGID-TASC scores are shown as mean (±SE).

Table 2. Rigid bronchoscopy tool for assessment of skills and competence scores of subcategories of rigid bronchoscopy intubation and navigation

Subcategories of Rigid Bronchoscopy	Novice	Intermediate	Expert	*P* Value for Difference
Preintubation score	12.5 ± 0.7	13.1 ± 1.3	17.2 ± 0.5	0.002
Upper airway navigation score	14 ± 2.3	22.7 ± 1.1	27.2 ± 0.3	<0.001
Lower airway navigation score	18.9 ± 1.1	24.7 ± 0.9	28.1 ± 0.5	<0.001
Attempts score	12 ± 1.1	14.5 ± 0.5	14.8 ± 0.2	0.02
Score for time required to complete	0.8 ± 0.8	3.2 ± 1.3	7.2 ± 0.8	<0.001

bronchoscope assembly and positioning of patient, upper and lower airway navigation, number of attempts, and time required to complete the procedure.

An important aspect of RIGID-TASC is an emphasis on the time required to complete the procedure (from introduction of bronchoscope into the mouth to manipulation up to carina in less than 60 s) and the number of attempts required to do so. Although a "rushed" procedure is never desirable, the interval from scope introduction to full intubation is the period in which the patient has undergone anesthesia induction and is thus apneic or hypopneic, while lacking a secure airway. Thus, the duration of this component of the procedure is indeed critical. This is especially true in light of the fact that many patients undergoing rigid bronchoscopy have multiple comorbidities, particularly critical central airway obstruction. Hence, a safe and atraumatic yet rapid intubation is integral to the overall principles of safety and effectiveness.

In a survey of pulmonary and critical care fellowship program directors, Pastis and colleagues (2) demonstrated that only 18% of the surveyed programs offered formal training in rigid bronchoscopy (range of 0–200 procedures). As expected, programs with a dedicated interventional pulmonology service had a higher likelihood of exposure to rigid bronchoscopy. However, the majority of trainees did not achieve exposure to the threshold number of procedures considered optimal for basic competence. In a recent survey by Yarmus and colleagues (3) of interventional pulmonology fellowship programs, the mean number of rigid bronchoscopies reported by the fellows was 122 (range, 62–187), clearly demonstrating that a dedicated year of training can achieve a higher number of procedures. However, it remains unclear as to what criteria can best be used to assess this competence.

Based on the numeric procedure-log criteria suggested by ACCP and the American Thoracic Society/European Respiratory Society, and procedural data from different surgical and procedural medical specialties, an arbitrary threshold number of procedures continues to be the prevalent way to assess procedural competence (4, 5, 8, 25–28). Different health care systems and institutions mostly rely on number of procedures to grant practicing privileges (29). However, it is clear that learners achieve different levels of procedural expertise at the same "threshold" number, as demonstrated by Wahidi and colleagues (7, 8). Hence, we submit that the incorporation of objective tools, such as RIGID-TASC, can help standardize the assessment of competency in procedural operators.

The Accreditation Council for Graduate Medical Education (ACGME) and the American Board of Medical Specialties have been actively engaged in developing objective methodologies to measure competence. There is a shift in the evaluation and accreditation of educational programs, away from education potential to outcomes and competence assessment, and the achievement of milestones (19, 30–33). Milestones have been defined as knowledge, skills, attitudes, and other attributes for each of the ACGME competencies that describe the development of competence from an early learner, up to and beyond that expected for unsupervised practice (33).

Competency-based education requires that each competency be not only teachable and learnable, but also measurable (34). As opposed to process-based education, the competency-based paradigm demands that the effectiveness of any educational intervention be assessed and validated by the use of objective metrics, devised specifically to measure the acquired skills. Competency has been defined in the Miller pyramid as the ability gained from knowledge and skills, and it forms a basis for performance

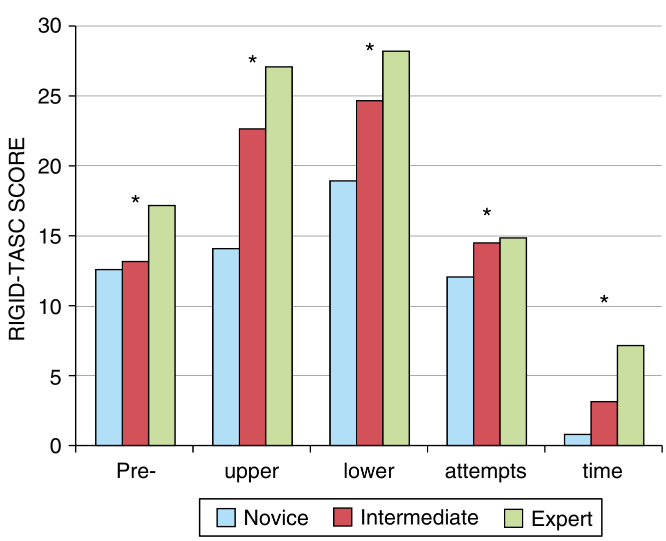

Figure 2. Rigid Bronchoscopy Tool for Assessment of Skills and Competence (RIGID-TASC) scores of subcategories of rigid bronchoscopy intubation. Higher score depicts better performance. *Statistically significant difference. attempts = attempts score; lower = lower airway navigation score; Pre- = preintubation score; time = score for time required to complete navigation up to carina; upper = upper airway navigation score. The RIGID-TASC scores are shown as mean (±SE).

(24). Thus, by measuring the operators' rigid bronchoscopy skills, we are in essence measuring the basis for their rigid bronchoscopy performance. Although ACGME does not currently accredit interventional pulmonology training programs, its recommendations can be adopted to improve interventional pulmonology educational outcomes. Development of tools, such as RIGID-TASC, can meet these goals (31).

An important finding that warrants keen attention is that the mean score of the intermediate group was 78.15, which is less than the experts' mean at 94.40, but was above the 75% mark. Is this good enough? We believe the answer is an emphatic "*no*." This relative ceiling effect is commonly seen in tests designed for "mastery testing" (35–39) (the use of the term "mastery" here should not be confused with its common usage, as in master of classical opera or master of renaissance literature). Mastery testing is a subtype of criterion-referenced testing, whereby we measure what proportion of a "total criterion sum" of knowledge and/or skills the learner has mastered, with the goal being as close as possible to the perfect 100%. This is the ideal testing paradigm for critical and high-risk procedures.

The rationale seems obvious: if a cardiac surgeon performs only 90% of the necessary steps of a coronary artery bypass grafting safely and effectively, the patient may never leave the surgical theater alive. Thus, in the mastery training and testing paradigm, the expected score is 100% or almost 100%. The "almost" is allowed for test-taking errors not reflected in clinical practice. High-risk and highly invasive procedures, such as rigid bronchoscopy, are ideal for mastery training and testing, because every operator must *master* each and every constituent component of a complex procedure, without exception, to achieve competence and be able to perform it safely and effectively. At the end of training, the variable distinguishing different trainees is not the perfect final score, but how long it took to get there (i.e., the slope of the curve and not the peak) (40–42).

High-quality assessment methods can be based on van der Vleuten and Shuwirth's utility equation: utility = validity × reliability × acceptability × educational impact × cost effectiveness (43). Validity refers to whether an instrument actually measures what it purports to measure (31, 43). RIGID-TASC was able to successfully discriminate individuals at different level of experience and expertise. Reliability refers to the reproducibility of the scores obtained from an assessment (31, 43, 44). RIGID-TASC is simple and clear, which ensures that different scorers will interpret and use its items similarly. The high interrater reliability of RIGID-TASC means that different independent examiners will place an operator at a certain level of skill in the same score category.

The acceptability of an assessment method is generally improved with its impact on the desired outcome, its ease of administration, and cost. The educational impact of a tool derives from its ability to stimulate self-awareness and positive change, to improve knowledge and skill acquisition in the learner (31, 43). Deconstructing rigid bronchoscopy into its main constituent steps can help improve the consistency of training. It can be used for immediate, detailed, and objective feedback to the learner, thus leading to deliberate practice to remedy deficiencies. Its use is simple for the instructor, hence disposing with the inherent heterogeneity and subjectivity of grading, thus facilitating its acceptance in a busy training program.

The need for objective competency-based assessment tools is further underscored by studies demonstrating the poor correlation between subjective faculty evaluations and standardized assessment scores (45, 46). Tools, such as RIGID-TASC, can be used by learners and instructors to verifiably track and document the gradual acquisition of knowledge and skills, as well as to identify areas where improvements are needed. Instructors and program directors can use the results to ascertain the strengths, weaknesses, and educational value of their curricula.

Limitations

The primary limitation of this study is that the assessment of skills was performed on a manikin. The same approach has been used in development of other valid assessment tools, such as the BSTAT (13). Even when using a rigid bronchoscope with integrated telescope (Richard Wolf Medical Instruments Corp., Vernon Hills, IL), or in patients with endotracheal tube in place, most of the assessment points of RIGID-TASC can be used.

Patient-based rigid bronchoscopy usually involves multiple complex therapeutic procedures, including tumor debridement, heat or cold therapy, and stent placement. Obviously, RIGID-TASC has not been designed to measure competence in these complex procedural skills. It is precisely designed to assess safe and effective rigid bronchoscopic intubation and central airway navigation. When learners in rigid bronchoscopy are objectively assessed as they progress along the spectrum, from novice to beginner to expert, a certain point is reached where each operator shows competence in the basic skills of rigid bronchoscopy: *safe and atraumatic positioning, intubation, and central airway navigation*. RIGID-TASC has been devised as an objective milestone for this level of competence. Training and assessment in all the other complex adjunct procedures (tumor resection, airway dilation, laser and other thermal treatments, stenting, etc.) will obviously need to follow acquisition of this basic level of competency.

Another limitation is that the rigid bronchoscope with 10-mm outer diameter was used for the testing, rather than the more commonly used 12-mm instrument. The 10-mm rigid bronchoscope was chosen, as the authors observed in pilot cases that most of the steps measured by the tool (such as assembly, patient positioning, protecting the lips and teeth, staying centered while navigating the airway, etc.) were not affected by the size of the bronchoscope. However, because the manikin airway (Nakhosteen Bronchoscopy Model) was originally designed for flexible bronchoscopy, passage of a 12-mm scope through the hard plastic vocal cords made rigid intubation of the manikin almost impossible for the novice and intermediate groups. However, the benefit of using the 10-mm bronchoscope was uniform to all three groups. In the real-life patient, of course, an experienced operator's skills will not show a marked difference with the size of the bronchoscope.

Conclusions

RIGID-TASC can accurately score the basic rigid bronchoscopy performance of operators based on their level of experience, with a high interrater reliability. It may be an effective instrument to provide objective feedback to rigid bronchoscopy trainees. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

1 Ayers ML, Beamis JF Jr. Rigid bronchoscopy in the twenty-first century. *Clin Chest Med* 2001;22:355–364.
2 Pastis NJ, Nietert PJ, Silvestri GA; American College of Chest Physicians Interventional Chest/Diagnostic Procedures Network Steering Committee. Variation in training for interventional pulmonary procedures among US pulmonary/critical care fellowships: a survey of fellowship directors. *Chest* 2005;127: 1614–1621.
3 Yarmus L, Feller-Kopman D, Imad M, Kim S, Lee HJ. Procedural volume and structure of interventional pulmonary fellowships: a survey of fellows and fellowship program directors. *Chest* 2013;144: 935–939.
4 Bolliger CT, Mathur PN, Beamis JF, Becker HD, Cavaliere S, Colt H, Diaz-Jimenez JP, Dumon JF, Edell E, Kovitz KL, *et al*.; European Respiratory Society/American Thoracic Society. ERS/ATS statement on interventional pulmonology. *Eur Respir J* 2002;19: 356–373.
5 Lamb CR, Feller-Kopman D, Ernst A, Simoff MJ, Sterman DH, Wahidi MM, Kovitz KL. An approach to interventional pulmonary fellowship training. *Chest* 2010;137:195–199.
6 Ernst A, Wahidi MM, Read CA, Buckley JD, Addrizzo-Harris DJ, Shah PL, Herth FJ, de Hoyos Parra A, Ornelas J, Yarmus L, *et al*. Adult bronchoscopy training: current state and suggestions for the future: CHEST expert panel report. *Chest* 2015;148:321–332.
7 Wahidi MM, Silvestri GA, Coakley RD, Ferguson JS, Shepherd RW, Moses L, Conforti J, Que LG, Anstrom KJ, McGuire F, *et al*. A prospective multicenter study of competency metrics and educational interventions in the learning of bronchoscopy among new pulmonary fellows. *Chest* 2010;137:1040–1049.
8 Wahidi MM, Hulett C, Pastis N, Shepherd RW, Shofer SL, Mahmood K, Lee H, Malhotra R, Moser B, Silvestri GA. Learning experience of linear endobronchial ultrasound among pulmonary trainees. *Chest* 2014;145:574–578.
9 Downing SM. Twelve steps for effective test development. In: Downing SM, Haladyna TM, editors. Handbook of test development. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers; 2006. pp. 3–25.
10 Downing SM, Tekian A, Yudkowsky R. Procedures for establishing defensible absolute passing scores on performance examinations in health professions education. *Teach Learn Med* 2006;18: 50–57.
11 Downing SM, Yudkowsky R. Introduction to assessment in the health professions. In: Downing SM, Yudkowsky R, editors. Assessment in health professions education. New York, NY: Routledge; 2009. pp. 1–20.
12 Yudkowsky R, Downing SM, Tekian A. Standard setting. In: Downing SM, Yudkowsky R, editors. Assessment in health professions education. New York, NY: Routledge; 2009. pp. 119–148.
13 Davoudi M, Osann K, Colt HG. Validation of two instruments to assess technical bronchoscopic skill using virtual reality simulation. *Respiration* 2008;76:92–101.
14 Davoudi M, Colt HG, Osann KE, Lamb CR, Mullon JJ. Endobronchial ultrasound skills and tasks assessment tool: assessing the validity evidence for a test of endobronchial ultrasound–guided transbronchial needle aspiration operator skill. *Am J Respir Crit Care Med* 2012;186:773–779.
15 Ernst A, Feller-Kopman D, Becker HD, Mehta AC. Central airway obstruction. *Am J Respir Crit Care Med* 2004;169: 1278–1297.
16 Dumon JF. A dedicated tracheobronchial stent. *Chest* 1990;97: 328–332.
17 Ishman SL, Brown DJ, Boss EF, Skinner ML, Tunkel DE, Stavinoha R, Lin SY. Development and pilot testing of an operative competency assessment tool for pediatric direct laryngoscopy and rigid bronchoscopy. *Laryngoscope* 2010;120:2294–2300.
18 Ishman SL, Benke JR, Johnson KE, Zur KB, Jacobs IN, Thorne MC, Brown DJ, Lin SY, Bhatti N, Deutsch ES. Blinded evaluation of interrater reliability of an operative competency assessment tool for direct laryngoscopy and rigid bronchoscopy. *Arch Otolaryngol Head Neck Surg* 2012;138:916–922.
19 Bancroft GN, Basu CB, Leong M, Mateo C, Hollier LH Jr, Stal S. Outcome-based residency education: teaching and evaluating the core competencies in plastic surgery. *Plast Reconstr Surg* 2008;121: 441e–448e.
20 Nunnally JC, Bernstein IH. Validity. In: Nunnally JC, Bernstein IH, editors. Psychometric theory, 3rd ed. New York, NY: McGraw-Hill; 1994. pp. 84–93.
21 Morris MC, Gallagher TK, Ridgway PF. Tools used to assess medical students competence in procedural skills at the end of a primary medical degree: a systematic review. *Med Educ Online* 2012;17: 10.3402/meo.v17i0.18398.
22 Haugen AS, Softeland E, Almeland SK, Sevdalis N, Vonen B, Eide GE, Nortvedt MW, Harthug S. Effect of the World Health Organization checklist on patient outcomes: a stepped wedge cluster randomized controlled trial. *Ann Surg* 2015;261:821-828.
23 Ahmed A, Ishman SL, Laeeq K, Bhatti NI. Assessment of improvement of trainee surgical skills in the operating room for tonsillectomy. *Laryngoscope* 2013;123:1639–1644.
24 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(9 suppl):S63–S67.
25 McGrath PD, Wennberg DE, Dickens JD Jr, Siewers AE, Lucas FL, Malenka DJ, Kellett MA Jr, Ryan TJ Jr. Relation between operator and hospital volume and outcomes following percutaneous coronary interventions in the era of the coronary stent. *JAMA* 2000;284: 3139–3144.
26 Hannan EL, Wu C, Walford G, King SB III, Holmes DR Jr, Ambrose JA, Sharma S, Katz S, Clark LT, Jones RH. Volume–outcome relationships for percutaneous coronary interventions in the stent era. *Circulation* 2005;112:1171–1179.
27 Torrington KG. Bronchoscopy training and competency: how many are enough? *Chest* 2000;118:572–573.
28 Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Ann Intern Med* 2002;137:511–520.
29 Tracy EE, Zephyrin LC, Rosman DA, Berkowitz LR. Credentialing based on surgical volume, physician workforce challenges, and patient access. *Obstet Gynecol* 2013;122:947–951.
30 Long DM. Competency-based residency training: the next advance in graduate medical education. *Acad Med* 2000;75:1178–1183.
31 Swing SR, Clyman SG, Holmboe ES, Williams RG. Advancing resident assessment in graduate medical education. *J Grad Med Educ* 2009; 1:278–286.
32 Swing SR, Beeson MS, Carraccio C, Coburn M, Iobst W, Selden NR, Stern PJ, Vydareny K. Educational milestone development in the first 7 specialties to enter the next accreditation system. *J Grad Med Educ* 2013;5:98–106.
33 Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system—rationale and benefits. *N Engl J Med* 2012; 366:1051–1056.
34 Brasel KJ, Bragg D, Simpson DE, Weigelt JA. Meeting the Accreditation Council for Graduate Medical Education competencies using established residency training program assessment tools. *Am J Surg* 2004;188:9–12.
35 Wayne DB, Butter J, Siddall VJ, Fudala MJ, Wade LD, Feinglass J, McGaghie WC. Mastery learning of advanced cardiac life support skills by internal medicine residents using simulation technology and deliberate practice. *J Gen Intern Med* 2006;21: 251–256.
36 Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 2004;79(10 suppl):S70–S81.
37 Zendejas B, Cook DA, Bingener J, Huebner M, Dunn WF, Sarr MG, Farley DR. Simulation-based mastery learning improves patient outcomes in laparoscopic inguinal hernia repair: a randomized controlled trial. *Ann Surg* 2011;254:502–509, discussion 509–511.
38 Popham WJ, Husek TR. Implications of criterion-referenced measurement. *J Educ Meas* 1969;6:1–9.
39 Wolcox RR. A note on the length and passing score of a mastery test. *J Educ Stat* 1976;1:359–364.

40 McGaghie WC, Siddall VJ, Mazmanian PE, Myers J; American College of Chest Physicians Health and Science Policy Committee. Lessons for continuing medical education from simulation research in undergraduate and graduate medical education: effectiveness of continuing medical education: American College of Chest Physicians Evidence-Based Educational Guidelines. *Chest* 2009;135(3 suppl): 62S–68S.
41 Yudkowsky R, Park YS, Lineberry M, Knox A, Ritter EM. Setting Mastery Learning Standards. *Acad Med* 2015;90:1495–500.
42 Lineberry M, Soo Park Y, Cook DA, Yudkowsky R. Making the case for mastery learning assessments: key issues in validation and justification. *Acad Med* 2015;90:1445–1450.
43 van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ* 2005;39:309–317.
44 Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ* 2004;38:1006–1012.
45 Elfenbein DM, Sippel RS, McDonald R, Watson T, Scarborough JE, Migaly J. Faculty evaluations of resident medical knowledge: can they be used to predict American Board of Surgery In-Training Examination performance? *Am J Surg* 2015;209:1095–1101.
46 Goldstein SD, Lindeman B, Colbert-Getz J, Arbella T, Dudas R, Lidor A, Sacks B. Faculty and resident evaluations of medical students on a surgery clerkship correlate poorly with standardized exam scores. *Am J Surg* 2014;207:231–235.

APPENDIX

RIGID BRONCHOSCOPY TOOL FOR ASSESSMENT OF SAFETY AND COMPETENCE

RIGID-TASC

*PARTICIPANT ID:*___________________________

*NO. OF RIGID BRONCHOSCOPY:*____________

*DATE OF PROCEDURE:*____________________

Pre-procedure Assessment:

1. Assembles the bronchoscope correctly [3]
   a. Yes Score 1
   b. No Score 0

Intra-procedure Assessment:

2. Positions the head of the patient correctly (hyperextended and lower than the body) [3]
   a. Yes Score 1
   b. No Score 0

3. Protects the upper and lower lips and teeth with left (non-dominant) hand during the introduction of the bronchoscope into the upper airway [2]
   a. Yes Score 1
   b. No Score 0

4. Holds the bronchoscope with right (dominant) hand [2]
   a. Yes Score 1
   b. No Score 0

5. Holds the bronchoscope with bevel up [2]
   a. Yes Score 1
   b. No Score 0

6. Stablizes the telescope in the rigid bronchoscope barrel with the thumb and avoids unnecessary movement of the telescope [3]
   a. Yes Score 1
   b. No Score 0

7. Positions the telescope in the rigid bronchoscope barrel so that the whole distal rim of the rigid bronchoscope is visible [3]

   a. Yes   Score 1
   b. No   Score 0

8. Positions the rigid suction in the rigid bronchoscope barrel so that it's neither protruding out/applying too much suction, nor too proximal to provide inadequate suction [3]
   a. Yes   Score 1
   b. No   Score 0

9. Introduces the rigid bronchoscope at ~90* angle into the upper airway and slowly increases the angle to 180* (parallel) to the airway while advancing the scope [4]
   a. Yes   Score 1
   b. No   Score 0

10. Follows the tongue/ uvula to find the epiglottis [4]
   a. Yes   Score 1
   b. No   Score 0

11. Elevates the epiglottis with the rigid bronchoscope [4]
   a. Yes   Score 1
   b. No   Score 0

12. Able to obtain full view of larynx/ vocal cords [4]
   a. Yes   Score 1
   b. No   Score 0

13. Avoids contact/injury to the arytenoids [4]
   a. Yes   Score 1
   b. No   Score 0

14. Able to insert the bronchoscope through the vocal cords with 90* rotation so that the bevel is parallel to the vocal cords [4]
   a. Yes   Score 1
   b. No   Score 0

15. Advances the rigid scope further with bevel down in the trachea [4]
   a. Yes   Score 1
   b. No   Score 0

16. Keeps the rigid scope centered in the airway lumen and avoids contact of the distal tip with the posterior trachea/ airway wall to minimize trauma [5]
   a. Yes   Score 1 (No episodes of contact/trauma)
   b. No   Score 0 (: 1 episodes of contact/trauma)

17. No. of attempts required to insert the bronchoscope into the trachea:

1 attempt:	Score 3 [5]
2 attempts:	Score 2 [5]
3 attempts:	Score 1 [5]
Unsuccessful or >3 attempts:	Score 0

18. Positions the upper teeth guard/protection correctly [3]
    a. Yes Score 1
    b. No Score 0

Advanced intra-procedural skills:

19. Turns the head to the *opposite* side to advance the rigid scope into the right mainstem bronchus [5]
    a. Yes Score 1
    b. No Score 0

20. Able to manipulate the rigid scope into the distal bronchus intermedius [5]
    a. Yes Score 1
    b. No Score 0

21. Turns the head to the *opposite* side to advance the rigid scope into the left mainstem bronchus [5]
    a. Yes Score 1
    b. No Score 0

22. Able to manipulate the rigid scope into the distal left mainstem bronchus [5]
    a. Yes Score 1
    b. No Score 0

23. Able to complete the procedure from introduction of bronchoscope into the mouth to manipulation up to carina in <60 sec. [8]
    a. Yes Score 1
    b. No Score 0

Time required to perform rigid intubation: ___________sec

The numbers in brackets [XX] should be multiplied with the score of each point (0, 1, 2, or 3) to determine the total score for that item.