Validation of an Endobronchial Ultrasound Simulator: Differentiating Operator Skill Level

David R. Stather a Paul MacEachern a Karen Rimmer a Christopher A. Hergott b Alain Tremblay a

a Division of Respiratory Medicine, University of Calgary, Calgary, Alta., and b Division of Respirology, Critical Care and Sleep Medicine, University of Saskatchewan, Saskatoon, Sask., Canada

Key Words
Bronchoscopy · Education · Endobronchial ultrasound · Interventional bronchoscopy · Simulation · Training

Abstract
Background: Endobronchial ultrasound (EBUS) is a revolutionary diagnostic procedure. There is currently no accepted method of assessing EBUS technical skill or competency.

Objectives: This study aimed to validate a computer EBUS simulator in differentiating between operators of varying clinical EBUS experience.

Methods: A convenience sample (n = 22) of bronchoscopists was separated into four cohorts based on previous bronchoscopy experience: group A = novice bronchoscopists, no EBUS experience (n = 4), group B = expert bronchoscopists, no EBUS experience (n = 5), group C = basic clinical EBUS training (n = 9), group D = EBUS experts (n = 4). After a standardized introduction session on the EBUS simulator, participants performed 2 simulated cases on an EBUS simulator with performance metrics measured by the simulator.

Results: Significant differences between groups were noted for total procedure time, percentage of lymph nodes identified (p < 0.05). Group D performed significantly better than all other groups for total procedure time and percentage of lymph nodes identified (p < 0.05). Group C performed significantly better than groups A and B for total procedure time, percentage of lymph nodes identified and percentage of successful biopsies (p < 0.05, ANOVA).

Conclusions: An EBUS simulator can accurately discriminate between operators with different levels of clinical EBUS experience. EBUS simulators show promise as a tool for assessing training and evaluating competency.

Introduction

Assessing competency in medical procedures is vital to patient care in order to ensure that an adequate pool of physicians competent in medical procedures exists to provide these important services. The assessment of procedural skills is also important in evaluating trainees' progress during fellowship programs, in assessing learn-
Endobronchial ultrasound (EBUS) is a minimally invasive pulmonary procedure used to diagnose pulmonary diseases by allowing visualization and safe, highly accurate biopsies of intrathoracic structures during bronchoscopy [7]. The American College of Chest Physicians guidelines for interventional pulmonary procedures state: ‘EBUS requires intensified training and practical experience in interpreting sonographic images, since the anatomic structures of the mediastinum are comparatively complex’ [8]. The European Respiratory Society/American Thoracic Society Taskforce recommends that trainees should be supervised for 40 EBUS procedures and should perform 25 procedures per year to maintain competency [9], although it is important to note that these recommendations are based on expert opinion.

Simulation training has been used for decades in the aviation and aerospace industries, where technical skills are learned and competency is assessed through simulation before learners are placed in real-life high-stakes situations. The use of simulation-based training for medical and surgical procedures has been demonstrated to be effective, cost-effective, and to increase patients’ comfort and safety [10–14]. Procedural simulators have been validated for assessing a variety of medical and surgical procedural skills including colonoscopy and laparoscopy [15, 16]. As of 2010, successful completion of courses utilizing simulation training (Fundamentals of Laparoscopic Surgery and Advanced Trauma Life Support) are now required for certification by the American Board of Surgery [17]. Demonstrating skill transfer from the actual to the simulation environment is necessary if simulators are to be used to measure procedural skill. With a valid task simulator the simulation should mimic the actual activity, such that experts should outperform novices. In addition to having been shown to be a useful tool for enhancing learning in basic bronchoscopy [18, 19], the use of a bronchoscopy simulator to measure basic bronchoscopy skill of learners has also been previously validated [20]. Similar studies have not been performed using an EBUS-specific simulator. This study aimed to validate a computer EBUS simulator in differentiating between operators of varying clinical EBUS experience.

Materials and Methods

Study Design
A prospective, nonrandomized comparative study of four groups of subjects was carried out:

Group A: novice bronchoscopists – junior pulmonary medicine trainees with >25 flexible bronchoscopy procedures experience, >4 months of pulmonary fellowship training and no EBUS experience.

Group B: expert bronchoscopists but novice EBUS operators – attending pulmonary physicians with >500 flexible bronchoscopy procedures experience, >5 years clinical practice, who perform standard transbronchial needle aspiration, and no EBUS experience.

Group C: basic clinical EBUS training – senior pulmonary medicine trainees or recent graduates (within 1 year) with >75 flexible bronchoscopy procedures experience and completed 1-month elective with the interventional pulmonary medicine service with ≥15 and ≤25 EBUS procedures experience.

Group D: expert EBUS operators – interventional pulmonologists or thoracic surgeons with >200 EBUS procedures experience.

There is currently no accepted definition of an expert bronchoscopist. In a study validating a basic flexible bronchoscopy simulator [20], the expert bronchoscopists had >500 procedures experience.

Similarly, there is currently no accepted definition of an expert EBUS operator. At the University of Calgary, the interventional pulmonary medicine fellow is considered an expert EBUS operator after completing the year-long fellowship, during which they complete >200 EBUS procedures.

Subjects received a 30-min introduction session to the EBUS simulator, and one practice case to familiarize them with the equipment. Before the practice case, subjects underwent a standardized education process to ensure they knew the different lymph node stations and what is considered to be a complete EBUS lymph node examination. In order to simplify and standardize the examinations, every case was assumed to be for lung cancer staging with a left-sided tumor, and only the 11R, 4L, 7, 4L and 11L stations were to be examined in all patients. Subjects were not allowed to perform the practice case until they demonstrated a complete understanding of a correct lymph node station examination and the associated airway and vascular anatomy by correctly answering a series of standardized questions (Appendix A). All subjects were then given the same two EBUS cases to complete with EBUS performance metrics recorded by the EBUS simulator.

EBUS Simulator Description
The EBUS simulator used for this study is the AccuTouch Flexible Bronchoscopy Simulator (CAE Healthcare, Montreal, Que., Canada), equipped with an EBUS module. This simulator has
Validation of an EBUS Simulator: Differentiating Operator Skill Level

been described in detail previously [20, 21]. The EBUS module consists of EBUS software and an EBUS needle system, but all other hardware components of the simulator remain the same (fig. 1, 2). At the time of completion of the manuscript there were no published studies utilizing such an EBUS simulator.

Measurement Tools
Subjects were given a pre-simulation questionnaire which assessed demographic information including computer, video game and bronchoscopy experience (Appendix B).

The EBUS simulator software comes with built-in metrics that are measured during each simulated case. These include: total procedure time, time to intubation, number of intubation attempts, number of times bronchoscope collided with closed vocal cords, number of vessel punctures, number of times needle was removed while suction was still on (contamination), number of times needle passed through far side of node, number of times needle advanced in bronchoscope without the correct use of the protective sheath (bronchoscope damage), quantity of lidocaine used, quantity of sedation used and vocal cord trauma. Two additional clinically relevant measurement metrics included were the percentage of lymph nodes correctly identified during lymph node ultrasound examination and the percentage of lymph nodes successfully biopsied.

Statistical Analysis
Three performance metrics were selected a priori as most relevant to the assessment of basic EBUS skill based on the authors’ experience in performing and teaching EBUS (primary outcome measures): total procedure time, percentage of lymph nodes correctly identified on lymph node examination and percentage of successful lymph node biopsies [successful lymph node biopsies/(failed lymph node biopsies + successful lymph node biopsies)].

The results of all metrics were analyzed between groups using the averaged performance from the two EBUS cases, with ANOVA comparison of means and post hoc analysis between individ-
ual groups with appropriate corrections for multiple comparisons with the least significant difference test.

Our sample size was based on convenience and availability of bronchoscopists of each skill level to participate in the study. Nevertheless, using a standard t test and including a minimum of 4 subjects per group should allow the detection of a 25% difference in the percentage of lymph nodes correctly identified on EBUS examination between groups (α of 0.05 and power of 0.80 with a standard deviation of 10%).

The study has received approval by the Calgary Health Research Ethics Board (ethics ID No. 22715) and all subjects gave written informed consent.

**Results**

Twenty-two subjects from three centers were enrolled in the study: group A (n = 4), group B (n = 5), group C (n = 9) and group D (n = 4). All subjects were able to complete the study. Table 1 lists the demographic data of the subjects.

The outcome measures as they relate to EBUS and bronchoscopy experience level for the combined EBUS simulator test cases are shown in table 2. Significant differences between groups were noted for the 3 primary outcome measures of total procedure time, percentage of lymph nodes correctly identified on EBUS examination and percentage of successful biopsies (p < 0.05, ANOVA; table 2; fig. 3). Mean total procedure times for groups A, B, C and D were 34.67, 28.80, 20.11 and 14.96 min, respectively. Mean percentages of lymph nodes identified for groups A, B, C and D were 47.8, 40.8, 67.1 and 91.1%, respectively. Mean percentages of successful biopsies for groups A, B, C and D were 64.2, 63.9, 91.1 and 100%, respectively. Comparisons between individual groups were performed with statistics reported in table 2.

Significant differences between groups were also noted in the secondary metrics of number of intubation attempts, number of collisions with closed vocal cords, number of contaminated samples, number of penetrations beyond far side of lymph node and number of completely missed lymph nodes (table 2). There was no sig-

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group A novice bronchoscopist (n = 4)</th>
<th>Group B expert bronchoscopist (n = 5)</th>
<th>Group C basic clinical EBUS (n = 9)</th>
<th>Group D expert EBUS (n = 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age, years*</td>
<td>32.8 (29–37)</td>
<td>48.4 (37–58)</td>
<td>31.0 (28–36)</td>
<td>38.8 (33–48)</td>
</tr>
<tr>
<td>Women*</td>
<td>3 (75%)</td>
<td>1 (20%)</td>
<td>2 (22%)</td>
<td>0</td>
</tr>
<tr>
<td>Mean number of flexible bronchoscopy exams*</td>
<td>28.3 (25–33)</td>
<td>740 (500–1,500)</td>
<td>115.7 (75–175)</td>
<td>932.3 (500–2,500)</td>
</tr>
<tr>
<td>Mean number of EBUS exams*</td>
<td>0</td>
<td>0</td>
<td>17.8 (16–22)</td>
<td>391.3 (250–600)</td>
</tr>
<tr>
<td>Mean prior bronchoscopy simulator experience, min</td>
<td>0</td>
<td>0</td>
<td>7.8 (0–60)</td>
<td>20.0 (0–60)</td>
</tr>
<tr>
<td>Mean prior EBUS simulator experience, min*</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>2.9 (0–5)</td>
</tr>
<tr>
<td>Training</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IP attending</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Pulmonary attending</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Thoracics attending</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Pulmonary fellow</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Thoracics fellow</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Computer experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very experienced</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Somewhat experienced</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Inexperienced</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Video game usage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often (few times/week)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Occasional (few times/month)</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Seldom (few times/year)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Never</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Figures in parentheses are ranges, unless indicated otherwise. * p < 0.05, ANOVA.
significant difference in time to intubation between any of the groups, and no vascular punctures or bronchoscope damage in any of the groups. There were no significant differences between groups for the amount of lidocaine or sedation used and the only procedural complication was one case of vocal cord trauma by one of the novice bronchoscopists (group A).

Computer experience and video game experience were not significantly related to the outcome measures. There were no significant differences in computer experience and video game experience between groups.

**Discussion**

This is the first study to validate the use of an EBUS simulator in the assessment of EBUS skill level. Significant differences between groups of various clinical EBUS experience were noted for all of the three primary outcome measures selected. The expert EBUS operators outperformed all other groups for all primary outcome measures, except for percentage of successful biopsies where there was no statistically significant difference when compared to the basic clinical EBUS group. The basic clinical EBUS group also outperformed the novice and expert bronchoscopist groups in all primary outcome measures. Interestingly, the only significant difference noted between the novice and expert bronchoscopists (both groups with no clinical EBUS experience) was a shorter procedure time for the expert bronchoscopists, suggesting that EBUS is a skill separate from that of basic flexible bronchoscopy. Not only can the EBUS simulator differentiate between expert EBUS operators and bronchoscopists with no EBUS experience, it can also differentiate between expert EBUS operators and those with more basic clinical EBUS experience. It would appear that the skills acquired during clinical EBUS training and

<table>
<thead>
<tr>
<th>EBUS performance metric</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
</tr>
</thead>
<tbody>
<tr>
<td>(mean values per case)</td>
<td>novice bronchoscopist</td>
<td>expert bronchoscopist</td>
<td>basic clinical EBUS</td>
<td>expert EBUS</td>
</tr>
<tr>
<td>Mean total procedure time, min*</td>
<td>34.67 (2.31)b–d</td>
<td>28.80 (5.42)a, c, d</td>
<td>20.11 (3.45)a, b, d</td>
<td>14.96 (2.69)a–c</td>
</tr>
<tr>
<td>Mean lymph nodes identified, %*</td>
<td>47.8 (7.4)b–d</td>
<td>40.8 (7.0)c, d</td>
<td>67.1 (10.4)a, b, d</td>
<td>97.1 (3.7)a–c</td>
</tr>
<tr>
<td>Mean successful biopsies, %*</td>
<td>64.2 (11.2)b–d</td>
<td>63.9 (12.1)a, c, d</td>
<td>91.1 (8.2)a, b</td>
<td>100.0 (0.0)a, b</td>
</tr>
<tr>
<td>Time to intubation, min</td>
<td>1.93 (0.44)</td>
<td>1.71 (0.80)</td>
<td>1.41 (0.71)</td>
<td>1.00 (0.27)</td>
</tr>
<tr>
<td>Intubation attempts*</td>
<td>4.5 (1.6)d</td>
<td>5.0 (3.3)c, d</td>
<td>2.4 (1.8)b</td>
<td>1.0 (0.0)a, b</td>
</tr>
<tr>
<td>Collisions with closed vocal cords*</td>
<td>3.5 (1.6)d</td>
<td>4.0 (3.3)c, d</td>
<td>1.4 (1.8)b</td>
<td>0.0 (0.0)a, b</td>
</tr>
<tr>
<td>Contaminated samples*</td>
<td>0.75 (1.0)</td>
<td>2.00 (1.6)c, d</td>
<td>0.22 (0.4)b</td>
<td>0.00 (0.0)b</td>
</tr>
<tr>
<td>Penetrating far side of lymph node*</td>
<td>2.25 (0.5)c, d</td>
<td>1.40 (1.3)d</td>
<td>0.78 (0.8)b</td>
<td>0.00 (0.0)a, b</td>
</tr>
<tr>
<td>Completely missing lymph node*</td>
<td>2.25 (1.5)c, d</td>
<td>2.00 (1.9)c, d</td>
<td>0.22 (0.4)a, b</td>
<td>0.00 (0.0)a, b</td>
</tr>
</tbody>
</table>

Figures in parentheses are SD. * p < 0.05, ANOVA. a p < 0.05 versus A; b p < 0.05 versus B; c p < 0.05 versus C; d p < 0.05 versus D.
practice are directly transferable and measurable by the EBUS simulator.

Prior studies of bronchoscopy simulation have demonstrated the utility of using inanimate models combined with didactic teaching, animal models and computer simulation [19, 23–25]. A significant limitation of inanimate models is a lack of realism. The costs, ethical concerns and anatomical differences between species are significant limitations with the use of animal models. Bronchoscopy simulators have been shown to improve the establishment of fundamental bronchoscopy skills and improve airway anatomy knowledge [22, 26]. Other studies have provided evidence that bronchoscopy simulator training accelerates the acquisition of basic bronchoscopy skills and that performance-based competency metrics can be used to evaluate basic bronchoscopy skills [18, 20, 27]. Hospital credentialing committees want standards that can be used in determining privileges, and requiring a minimum number of procedures is a frequently used method [2, 6]. However, the American Board of Internal Medicine has not adopted this approach, presumably because of the lack of standardization of the procedures being performed, the significant variability in the rate of skill acquisition and the mounting evidence that the use of procedure numbers for establishing competency is inadequate [13, 18, 28, 29]. Recommendations for numbers of proctored EBUS procedures that need to be performed to demonstrate competency have been made [8, 9], but some of these relate to radial EBUS rather than the linear technique, and the recommendations were based on expert opinion only, without supporting evidence to justify their validity. Through the use of EBUS simulators we may now have a method to more objectively standardize the evaluation of competency for this technique.

There are some limitations to this study. Due to the limitations of a yearly fellowship program, there were small sample sizes in the groups. Confirmation of our findings with larger trials would be ideal. Although the metrics used were objective measurements, most of which were recorded directly by the simulator, the use of a single proctor introduces the possibility of a bias. A very specific, scripted protocol was followed for each subject during both the introduction to EBUS phase and during the evaluated cases to ensure that this potential for bias was minimized. Proctor EBUS teaching/assisting skill could have improved with time, therefore introducing bias; however, the majority of group C and D evaluations were completed before those of groups A and B. This potential bias would therefore have favored better performance in groups A and B, and complete elimination of the bias would most likely strengthen the observed difference between groups demonstrated in this study.

Although we selected the primary outcome measures used in this study (procedure time, percentage of lymph nodes correctly identified, percentage of lymph nodes successfully biopsied) based on our experience with EBUS education and believe them to be meaningful measures of EBUS performance, other metrics may be as or more important. Our primary metrics can assess the basic goal of the EBUS procedure: correctly identifying and sampling mediastinal lymph nodes. Procedural accuracy is critical to performing a quality EBUS procedure. When performing EBUS procedures for lung cancer staging, missing an important lymph node or incorrectly identifying a lymph node station has the potential to incorrectly upstage or downstage a patient, thus negatively impacting patient management. Our third primary metric of procedure time adds a measure of efficiency and safety for the procedure given that longer procedures may result in the administration of more sedation with its inherent risks, as well as increasing the likelihood that the patient will not tolerate procedure completion. While many of the additional metrics potentially representing important measures of EBUS skill were also statistically different between groups, the frequency of these events were relatively low and may not be as useful to assess an individual operator. Similarly, frank complications such as vocal cord injury and vascular puncture were very low in all groups.

While the simulator was able to differentiate between the four groups, this study does not in itself determine the minimal skill level threshold associated with the ability to safely, accurately and independently perform a clinical EBUS procedure. Further studies correlating performance on the simulator with performance during clinical cases will be required. It should also be emphasized that this study investigates the technical aspects of EBUS procedural skill and does not attempt to study the cognitive aspects of EBUS. Conclusions regarding cognitive aspects of EBUS, such as pre-procedure preparation of the patient, selection of appropriate lymph node targets and interpretation of results, cannot be made based on our data.

A number of problems were identified with the EBUS simulator, and with all types of simulation training there can be issues with realism and transfer. The white light image provided by the simulator while the EBUS bronchoscope is above the vocal cords is closer to that of a regular flexible bronchoscope, as opposed to the 30-de-
gree angle of view of the EBUS bronchoscope. The problem is corrected once the bronchoscope passes the vocal cords, but could confuse learners when learning the EBUS intubation technique. Lymph nodes are generally more easily identified than during clinical cases, with more distinct separation between nodes and vessels. A number of other software-related discrepancies were noted which did distract the intermediate and expert groups, such as the white light image not always corresponding correctly with the EBUS image (i.e. there may be an EBUS image when there is clearly no apposition of the transducer against the airway wall on the white light image). Finally, the EBUS simulator, including the needle and bronchoscope system, is fragile, resource intensive and expensive, with a current cost of more than USD 100,000 which may limit its widespread use.

In summary, our study is the first to demonstrate that EBUS simulators can accurately discriminate between operators with different levels of clinical EBUS experience. This suggests that there is significant skill transfer from clinical EBUS to the simulator. Demonstrating skill transfer from reality to the simulated environment does not, however, imply that there is skill transfer from the simulated environment to reality, that is, that learning on a simulator translates to improved clinical skills. Given these encouraging findings, EBUS simulators may prove to be useful tools in programs where assessment of EBUS technical competency is of interest such as in fellowship training programs, for the hospital procedure credentialing process and in continuing medical education sessions.

Financial Disclosure and Conflicts of Interest

D.R.S., P.M., C.A.H. and K.R. have no conflicts of interest to disclose. A.T. has received consulting fees from Olympus America. The University of Calgary has received grants from Olympus Canada for support of an Interventional Pulmonary Medicine Training Program and for continuing medical education events relating to EBUS.

Appendix A

Validation of an EBUS Simulator in Differentiating Operator Skill Level

Learner EBUS knowledge review protocol
- Name the 5 lymph node stations that you will examine: 11R, 4R, 7, 4L, 11L.
- Assuming a left-sided non-small cell lung cancer, list the N3 to N1 lymph node stations: 11R, 4R = N3, 7, 4L = N2 and 11L = N1.
- Assuming a right-sided non-small cell lung cancer, list the N3 to N1 lymph node stations: 11L, 4L = N3, 7, 4R = N2 and 11R = N1.
- Draw a diagram illustrating all 5 lymph node stations that you will examine, and a detailed relationship between the lymph node stations and all of the surrounding airways and landmark blood vessels (must repeat this drawing until they correctly identify all 5 stations including the correct relationship with the airways and the azygous vein, superior vena cava, inominate artery, aorta and left main pulmonary artery).

Appendix B

Validation of an EBUS Simulator in Differentiating Operator Skill Level

Study participant experience level
Participant study code: __________________________
Participant name: __________________________
Age: __________________________
Sex (circle one): Male Female
Computer experience (check one):
- never used a computer
- inexperienced
- somewhat experienced
- very experienced
Level of video game play:
- very often (daily)
- often (few times per week)
- occasional (few times per month)
- seldom (few times per year)
- never
Level of training (please circle one): PGY1, PGY2, PGY3, PGY4, PGY5, PGY6, PGY7, PGY8, IPM attending, thoracic surgery attending
Training program: __________________________
Number of bronchoscopies performed: __________________________
Number of EBUS performed: __________________________
Prior bronchoscopy simulator experience (min/h): __________________________
Prior EBUS simulator experience (min/h): __________________________

Validation of an EBUS Simulator: Differentiating Operator Skill Level

Respiration 2011;81:325–332

331
References


