Mastery Learning of Thoracentesis Skills by Internal Medicine Residents Using Simulation Technology and Deliberate Practice

Diane B. Wayne, MD1
Jeffrey H. Barsuk, MD1
Kevin J. O’Leary, MD1
Monica J. Fudala, BA1
William C. McGaghie, PhD2

1 Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois
2 Augusta Webster, MD, Office of Medical Education and Faculty Development, Northwestern University Feinberg School of Medicine, Chicago, Illinois

Supported by Northwestern Memorial Hospital.

The authors thank Charles Watts, MD, and J. Larry Jameson, MD, PhD, for their support of this work. We recognize and appreciate the Northwestern University internal medicine residents for their dedication to patient care and education.

BACKGROUND: Thoracentesis has been identified as a core competency for hospitalists. Residency training may not provide adequate preparation to perform this procedure.

OBJECTIVE: Our aim was to use a thoracentesis simulation to assess graduating residents’ proficiency in thoracentesis procedural skills and to evaluate the impact of an educational intervention on skill development to mastery standards.

DESIGN: This was a pretest–posttest mastery learning design without a control group.

SETTING: Participants were 40 third-year internal medicine residents in a university-affiliated program.

INTERVENTION: Using an observational checklist, each resident underwent baseline assessment of thoracentesis skill using a standardized clinical history, radiograph, and thoracentesis simulation. After baseline testing, residents received two 2-hour education sessions featuring a videotaped presentation and deliberate practice with the thoracentesis simulator.

MEASUREMENTS: Residents were retested after the intervention. Skill mastery was defined as meeting or exceeding the minimum passing score (MPS) set by an expert panel at thoracentesis posttest. Those who did not achieve the MPS had additional deliberate practice and were retested until the MPS was reached.

RESULTS: Performance improved 71% from pretest to posttest on the clinical skills examination. All residents met or exceeded the mastery standard. The amount of practice time needed to reach the MPS was the only predictor (negative) of posttest performance. The education program was rated highly.

CONCLUSIONS: A curriculum featuring deliberate practice dramatically increased the skills of residents in thoracentesis. Residents enjoy training and receiving evaluation and feedback in a simulated clinical environment. This mastery program illustrates a feasible and reliable mechanism to achieve procedural competency. Journal of Hospital Medicine 2008;3:48–54. © 2008 Society of Hospital Medicine.

KEYWORDS: thoracentesis, residency education, mastery learning, simulation-based education.

In a supplement to its inaugural issue, the Journal of Hospital Medicine published core competencies for hospitalists covering 3 areas: clinical conditions, systems in health care, and procedures.1 Completion of a traditional internal medicine residency may not provide hospitalists with the skills necessary to safely perform necessary procedures such as thoracentesis. A recent article reported that most internal medicine residents surveyed were uncomfortable performing common procedures, and their discomfort was higher for thoracentesis than for central line in-
sertion, lumbar puncture, or paracentesis. This confirmed a previous report that family practice residents had low confidence in performing thoracenteses. Thoracentesis also carries the risk of the potentially life-threatening complication of pneumothorax, which may be increased when performed by physicians-in-training.

One method for improving training and assessment is the use of simulation technology. Simulation has been used to increase knowledge, provide opportunities for deliberate and safe practice, and shape the development of clinical skills. Simulation has also been advocated for assessing competence in procedures including carotid angiography, emergency airway management, basic bronchoscopy, and advanced cardiac life support (ACLS).

Recently, we used simulation technology to help residents reach mastery learning standards for ACLS. Mastery learning, an extreme form of competency-based education, implies that learners have acquired the clinical knowledge and skill measured against rigorous achievement standards. In mastery learning, educational results are equivalent, whereas educational practice time differs. To demonstrate mastery learning, we first documented a 38% improvement in skill after a simulation-based educational intervention and used a multidisciplinary panel to determine mastery achievement standards for ACLS skills in 6 clinical scenarios. These standards were used in a study in which the amount of time needed to achieve skill mastery was allowed to vary while the skill outcomes of the residents were identical clinically.

The present study had 4 aims. The first was to assess the baseline skill and knowledge of third-year residents in thoracentesis. The second was to compare the thoracentesis-related knowledge and skills of residents before and after an educational intervention. The third was to assess the correlation of medical knowledge and clinical experience with performance on a clinical skills examination after simulation training. The last was to document the feasibility of incorporating simulation-based education into a training program.

METHODS
Objectives and Design
The study, which had a pretest–posttest design without a control group, was of a simulation-based, mastery learning educational intervention in thoracentesis. Primary measurements were obtained at baseline (pretest) and after the educational intervention (posttest).

Participants
Study participants were all 40 third-year residents in the internal medicine residency program at Northwestern University’s Chicago campus from January to May 2006. The Northwestern University Institutional Review Board approved the study. Participants provided informed consent before baseline assessment.

This residency program is based at Northwestern Memorial Hospital (NMH) and the Jesse Brown Veteran’s Affairs Medical Center. Residents perform thoracenteses under the supervision of second- or third-year residents or faculty members who are credentialed to perform the procedure. A didactic lecture on thoracentesis is part of the annual lecture series.

Procedure
The residents were kept as an intact group during the study period. The research procedure had 2 phases. First, the knowledge and clinical skills of participants at baseline were measured. Second, residents received two 2-hour education sessions featuring didactic content and deliberate practice using a thoracentesis model. Between 4 and 6 weeks after the pretest, all residents were retested and were expected to meet or exceed a minimum passing score (MPS) on the clinical skills exam. Those who scored below the MPS engaged in more clinical skills practice until the mastery standard was reached. The amount of extra time needed to achieve the MPS was documented.

Educational Intervention
The intervention was designed to help residents acquire the knowledge and skills needed to perform a competent thoracentesis. The necessary components for mastery skill development were contained in the intervention. These included deliberate practice, rigorous skills assessment, and the provision of feedback in a supportive environment.

The study was conducted in the Northwestern University Center for Advanced Surgical Education (N-CASE) using the thoracentesis simulator developed by MediSim Inc. (Alton, Ontario). The model features realistic skin texture, ribs, and a fluid filled reservoir. Needles of various sizes can be inserted and fluid withdrawn. The model also accommo-
dates the catheter/needle apparatus found in the thoracentesis kits (Tyco Healthcare®, Pembroke, Bermuda) used at NMH.

Teaching and testing sessions were standardized. In teaching sessions, groups of 2-4 residents had 4 hours to practice and ask questions, and to receive structured education and feedback from 1 of 2 hospitalist faculty instructors (J.H.B., K.J.O.). One of the 4 hours was devoted to the presentation of didactic material on indications, complications, and interpretation of results and a step-by-step demonstration of a thoracentesis. This presentation was videotaped to ensure standardization of content. The remaining 3 hours were devoted to clinical skills exam education, deliberate practice, and feedback.

One resident was present at each pretest and posttest session with 1 of the 2 faculty instructors who gave standardized instructions. The resident was expected to obtain a relevant history; perform a limited physical examination; review PA, lateral, and decubitus chest radiographs; perform a simulated thoracentesis; and order appropriate diagnostic tests. Written examinations were completed at the pretest and posttest sessions.

Measurements
A 25-item checklist was developed for the thoracentesis procedure using relevant sources and rigorous step-by-step procedures. Each skill or other action was listed in order and given equal weight. Each skill or action was scored dichotomously—either 0 = done correctly or 1 = done incorrectly. Checklists were reviewed for completeness and accuracy by 2 authors who frequently perform and supervise thoracenteses (J.H.B., K.J.O.), 2 authors with expertise in checklist design (D.B.W., W.C.M.), and the physician director of the medical intensive care unit at NMH. The checklist was used in a pilot clinical skills examination of 4 chief medical residents to estimate checklist reliability and face validity.

The MPS for the thoracentesis clinical skills examination was determined by 10 clinical experts using the Angoff and Hofstee standard setting methods. The panel was composed of clinical pulmonary critical care medicine faculty (n = 7) and senior fellows (n = 3). Each panel member was given instruction on standard setting and asked to use the Angoff and Hofstee methods to assign pass/fail standards. The Angoff method asks expert judges to estimate the percentage of “borderline” examinees who would answer each test item correctly. The Hofstee method requires judges to estimate 4 properties of an evaluation’s passing scores and failure rates. The panel was asked to repeat their judgments 6 weeks later to assure stability of the MPS. Details about the use of a standard setting exercise to set an MPS for clinical skills examinations have been published previously.

Evaluation of each resident’s skill was recorded on the checklist by 1 of the 2 faculty raters at the pretest and posttest sessions. A random sample of 50% of the pretest sessions was rescored by a third rater with expertise in scoring clinical skills examinations (D.B.W.) to assess interrater reliability. The rescorer was blinded to the results of the first evaluation.

A multiple choice written examination was prepared according to examination development guidelines using appropriate reference articles and texts. The examination was prepared by 1 author (J.H.B.) and reviewed for accuracy and clarity by 2 others (K.J.O., D.B.W.) and by the director of the medical intensive care unit at NMH. The examination had questions on knowledge and comprehension of the procedure as well as data interpretation and application. It was administered to 9 fourth-year medical students and 5 pulmonary/critical care fellows to obtain pilot data. Results of the pilot allowed creation of a pretest and a posttest that were equivalent in content and difficulty. The Kuder Richardson Formula 20 (KR-20) reliability coefficients for the 20-item pretest and the 20-item posttest were .72 and .74, respectively.

Demographic data were obtained from the participants including age, gender, ethnicity, medical school, and scores on the United States Medical Licensing Examination (USMLE) Steps 1 and 2. Each resident’s experience performing the procedure was also collected at pretest.

Primary outcome measures were performance on the posttest written and clinical examinations. Secondary outcome measures were the total training time needed to reach the MPS (minimum = 240 minutes) and a course evaluation questionnaire.

Data Analysis
Checklist score reliability was estimated by calculating interrater reliability, the preferred method for assessments that depend on human judges, using the kappa (k) coefficient adjusted using the formula of Brennan and Prediger. Within-group differences from pretest (baseline) to posttest (out-
come) were analyzed using paired \( t \) tests. Multiple regression analysis was used to assess the correlation of posttest performance on thoracentesis skills with (1) performance on pretest thoracentesis skills, (2) medical knowledge measured by the thoracentesis pretest and USMLE Steps 1 and 2, (3) clinical experience in performing thoracentesis, (4) clinical self-confidence about performing thoracentesis, and (5) whether additional training was needed to master the procedure.

**RESULTS**

All residents consented to participate and completed the entire training protocol. Table 1 presents demographic data about the residents. Most had limited experience performing and supervising thoracenteses.

Interrater reliability for the thoracentesis checklist data was calculated at pretest. Across the 25 checklist items, the mean kappa coefficient was very high \( (\kappa_{\text{m}} = .94) \). The MPS used as the mastery achievement standard was 80% (eg, 20 of 25 checklist items). This was the mean of the Angoff and Hofstee ratings obtained from the first judgment of the expert panel and is displayed in Figure 1.

No resident achieved mastery at pretest. However, 37 of the 40 medicine residents (93%) achieved mastery within the standard 4-hour thoracentesis curriculum. The remaining 3 residents (7%) needed extra time ranging from 20 to 90 minutes to reach mastery.

Figure 1 is a graphic portrait with descriptive statistics of the residents’ pretest and posttest performance on the thoracentesis written and clinical skills exams. For the written exam, the mean score rose from 57.63% to 89.75%, a statistically significant improvement of 56% from pretest to posttest \( (t_{39} = -17.0, P < .0001) \). The clinical skills exam also showed a highly significant 71% pretest-to-posttest gain, as the mean score rose from 51.70% to 88.3% \( (t_{39} = -15.6, P < .0001) \).

Results from the regression analysis indicate that neither pretest performance, medical knowledge measured by local or USMLE examinations, nor thoracentesis clinical experience was correlated with the posttest measure of thoracentesis clinical skills. However, the need for additional practice to reach the mastery standard on the posttest was a powerful negative predictor of posttest performance: \( \hat{b} = -.27 \) (95% CI = -.46 to -.09; \( P < .006; \hat{r}^2 = .28 \)). For those residents who required extra practice time, the initial clinical skills posttest score was 20% lower than that of their peers. Although the need for extra deliberate practice was associated with relatively lower initial posttest scores, all residents ultimately met or exceeded the rigorous thoracentesis MPS.

The responses of the 40 residents on a course evaluation questionnaire were uniformly positive. Responses were recorded on a Likert scale where 1 = strongly disagree, 2 = disagree, 3 = uncertain, 4 = agree, and 5 = strongly agree (Table 2). The data show that residents strongly agreed that practice with the medical simulator boosts clinical skills and self-confidence, that they received useful feedback from the training sessions, and that deliberate practice using the simulator is a valuable educational experience. Residents were uncertain whether practice with the medical simulator has more educational value than patient care.

**DISCUSSION**

This study demonstrates the use of a mastery learning model to develop the thoracentesis skills of internal medicine residents to a high level. Use of a thoracentesis model in a structured educational program offering an opportunity for deliberate

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PGY-3 Resident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>28.88 (1.57)</td>
</tr>
<tr>
<td>Male</td>
<td>23 (57.5%)</td>
</tr>
<tr>
<td>Female</td>
<td>17 (42.5%)</td>
</tr>
<tr>
<td>African American</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td>White</td>
<td>21 (52.5%)</td>
</tr>
<tr>
<td>Asian</td>
<td>14 (35.0%)</td>
</tr>
<tr>
<td>Other</td>
<td>4 (10.0%)</td>
</tr>
<tr>
<td>U.S. medical school graduate</td>
<td>39 (97.5%)</td>
</tr>
<tr>
<td>Foreign medical school graduate</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td>Number of thoracentesis procedures</td>
<td></td>
</tr>
<tr>
<td>Performed as an intern</td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>27.5%</td>
</tr>
<tr>
<td>2-4</td>
<td>60.0%</td>
</tr>
<tr>
<td>≥5</td>
<td>12.5%</td>
</tr>
<tr>
<td>Performed as a PGY-2 and PGY-3 resident</td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>25.0%</td>
</tr>
<tr>
<td>2-4</td>
<td>55.0%</td>
</tr>
<tr>
<td>≥5</td>
<td>20.0%</td>
</tr>
<tr>
<td>Supervised others as a PGY-2 and PGY-3 resident</td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>27.5%</td>
</tr>
<tr>
<td>2-4</td>
<td>57.5%</td>
</tr>
<tr>
<td>≥5</td>
<td>15.0%</td>
</tr>
</tbody>
</table>
practice with feedback produced large and consistent improvements in residents’ skills. An important finding of our study is that despite having completed most of their internal medicine training, residents displayed poor knowledge and clinical skill in thoracentesis procedures at baseline. This is similar to previous studies showing that the procedural skills and knowledge of physicians at all stages of training are often poor. Examples of areas in which significant gaps were found include basic skills such as chest radiography, emergency airway management, and pulmonary auscultation.

In contrast, after the mastery learning program, all the residents met or exceeded the MPS for the thoracentesis clinical procedure and scored much higher on the posttest written examination. Our data also demonstrate that medical knowledge measured by procedure-specific pretests and posttests and USMLE Steps 1 and 2 scores were not correlated with thoracentesis skill acquisition. This reinforces findings from our previous studies of ACLS skill acquisition and supports the difference between professional and academic achievement. Pretest skill performance and clinical experience also were not correlated with posttest outcomes. However, the amount of deliberate practice needed to reach the mastery standard was a powerful negative predictor of posttest thoracentesis skill scores, replicating our research on ACLS.

We believe that clinical experience was not correlated with posttest outcomes because residents infrequently performed thoracenteses procedures during their training.

**TABLE 2**
Course Evaluations Provided by All Residents (n = 40) after Simulation-Based Educational Program

<table>
<thead>
<tr>
<th>Practice Fact</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice with the thoracentesis model boosts my skills to perform this procedure.</td>
<td>4.3</td>
<td>0.8</td>
</tr>
<tr>
<td>I receive useful educational feedback from the training sessions.</td>
<td>4.0</td>
<td>0.6</td>
</tr>
<tr>
<td>Practice with the thoracentesis model boosts my clinical self-confidence.</td>
<td>4.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Practice with the thoracentesis model has more educational value than patient care experience.</td>
<td>2.3</td>
<td>1.0</td>
</tr>
<tr>
<td>The Skills Center staff are competent.</td>
<td>4.3</td>
<td>0.6</td>
</tr>
<tr>
<td>Practice sessions in the Skills Center are a good use of my time.</td>
<td>3.7</td>
<td>1.0</td>
</tr>
<tr>
<td>Practice sessions using procedural models should be a required component of residency education.</td>
<td>3.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Deliberate practice using models is a valuable educational experience.</td>
<td>4.0</td>
<td>0.9</td>
</tr>
<tr>
<td>Practice sessions using models are hard work.</td>
<td>2.1</td>
<td>0.7</td>
</tr>
<tr>
<td>Increasing the difficulty of simulated clinical problems helps me become a better doctor.</td>
<td>3.9</td>
<td>0.7</td>
</tr>
<tr>
<td>The controlled environment in the Skills Center helps me focus on clinical education problems.</td>
<td>3.9</td>
<td>0.8</td>
</tr>
<tr>
<td>Practice with the thoracentesis model has helped to prepare me to perform the procedure better than clinical experience alone.</td>
<td>4.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

**FIGURE 1.** Performance on thoracentesis written exam and clinical skills exam performance (MPS, minimum passing score).

![Graph showing performance on thoracentesis written exam and clinical skills exam.](image-url)
This project demonstrates a practical model for outcomes-based education, certification, and program accreditation. Given the need to move procedural training in internal medicine beyond such historical methods as “see one, do one, teach one,” extension of the mastery model to other invasive procedures deserves further study. At our institution we have been encouraged by the ability of simulation-based education in ACLS to promote long-term skill retention and improvement in the quality of actual patient care. In addition to studying these outcomes for thoracentesis, we plan to incorporate the use of ultrasound when training residents to perform procedures such as thoracentesis and central venous catheter insertion.

Given concerns about the quality of resident preparation to perform invasive procedures, programs such as this should be considered as part of the procedural certification process. As shown by our experience with several classes of residents (n = 158), use of simulation technology to reach high procedural skill levels is effective and feasible in internal medicine residency training. In addition, our residents have consistently enjoyed participating in the simulated training programs. Postcourse questionnaires show that residents agree that deliberate practice with simulation technology complements but does not replace patient care in graduate medical education.

An important question needing more research is whether performance in a simulated environment transfers to actual clinical settings. Several small studies have demonstrated such a relationship, yet the transfer of simulated training to clinical practice requires further study. More work should also be done to assess long-term retention of skills and to determine the utility and benefit of simulation-based training in procedural certification and credentialing.

This study had several limitations. It was conducted in 1 training program at a single medical center. The sample size (n = 40) was relatively small. The thoracentesis model was used for both education and testing, potentially confounding the events. However, these limitations do not diminish the pronounced impact that the simulation-based training had on the skills and knowledge of our residents.

In conclusion, this study has demonstrated the ability of deliberate practice using a thoracentesis model to produce high-level performance of simulated thoracenteses. The project received high ratings from learners and provides reliable assessments of procedural competence. Although internists are performing fewer invasive procedures now than in years past, procedural training is still an important component of internal medicine training. Attainment of high procedural skill levels may be especially important for residents who plan to practice hospital medicine. We believe that simulation-based training using deliberate practice should be a key contributor to future internal medicine residency education, certification, and accreditation.

Address for correspondence and reprint requests: Diane B. Wayne, MD, Department of Medicine, Northwestern University Feinberg School of Medicine, 251 E. Huron St. Galter 3-150, Chicago, IL 60611; Fax: (312) 926-6905; E-mail: dwayne@northwestern.edu

Received 23 March 2007; revision received 16 July 2007; accepted 22 July 2007.

REFERENCES


