# DEVELOPMENT AND TESTING OF A THORACOSTOMY ASSESSMENT TOOL THROUGH SELF, PEER AND EXPERT EVALUATION IN A SIMULATION ENVIRONMENT

By

MARISSA LOVETT

A Thesis Submitted to The Honors College

In Partial Fulfillment of the Bachelors Degree
With Honors in

Physiology

THE UNIVERSITY OF ARIZONA

MAY 2015

Approved by:

Dr. Allan Hamilton

Department of Surgery

DEVELOPMENT AND TESTING OF A THORACOSTOMY ASSESSMENT TOOL THROUGH SELF, PEER AND EXPERT EVALUATION IN A SIMULATION ENVIRONMENT

**Abstract**

*Background:* Self and peer evaluation using an objective assessment tool may be an effective way to evaluate procedural competency without requiring one-on-one expert attention in a simulation environment.

*Materials and Methods:* An assessment tool was created as a checklist of critical steps involved in performing a thoracostomy. The checklist was verified by three expert physicians. Sixteen inexperienced medical students were recruited in pairs to perform one thoracostomy on a synthetic chest tube model in a simulated environment. After completing the procedure, the students used the assessment tool alongside a video recording of their procedure to conduct a self and peer evaluation. An expert physician conducted an independent evaluation of each student using the same assessment tool. The variations between the self, peer, and expert evaluations were compared to test the assessment tool for objectivity.

*Results*: There was significant variation between self and expert evaluations (p = <0.00001 ) and between peer and expert responses (p = <0.00001). The significant variation occurred in 19 of the critical tasks

*Conclusion:* The assessment tool is not objective across the evaluators. Self and peer evaluation is an ineffective approach for assessing procedural competency, until variation is reduced among the 19 critical tasks.

## Introduction

Simulation education has become a standard in the field of medicine over the last two decades, allowing students to develop skills through repeatable and realistic simulations in a safe learning environment.[1] The ability to repeat procedures until an appropriate level of competence has been reached has contributed to the emergence of medical simulation as a replacement for the traditional "see one, do one, teach one" model of teaching.[1,2] Despite the many benefits of learning in a simulation environment, educators continue to debate how to effectively measure competency during simulation training.[3] Standardized approaches for assessing laparoscopic surgical skills have been adopted by the American Board of Surgery, but methods for evaluating specialized procedures have yet to be regulated.[1]

A primary barrier to the incorporation of simulation assessment in medical curricula is the requirement of individual feedback from an instructor.[1] Instructors are limited by their time availability, preventing individual feedback from being an efficient form of assessment.[2,4] Peer and self-evaluations have been used as an alternative to instructor feedback in laparoscopic surgical skill assessments and may provide a more efficient form of feedback and learning.[5,1]

An additional barrier to the progression of simulation education is the absence of a single, standardized assessment tool for specialized procedures.[4,6] Each procedure has a required list of steps that can be broken down into critical tasks and behaviors. An effective assessment tool needs to be adaptable to the different steps of each procedure, using a single format for each evaluation.[6] An ideal assessment tool would also need the ability to measure competence of each critical task using a clear evaluation scale.[6] The best level of specificity and simplicity within an evaluation scale is still under review, with the global rating scale remaining the most common.[3]

## Methods

### *Development of Assessment Protocol*

The assessment protocol was originally compiled from the New England Journal of Medicine Chest-Tube Insertion video and a variety of other emergency medicine thoracostomy protocols.[7,8,9,10,11] Repetitious steps were removed, and variations were filtered by available equipment. The procedural steps within the compiled thoracostomy protocol were then broken down into critical tasks. In this study, critical tasks are defined as single actions necessary to complete a step or part of a step of a procedure. To be an effective critical task, the action must be able to be visualized by the assessor and independent from other critical tasks. The critical tasks can be seen in the second column of the assessment tool in Figure 1.

| Emergent Surgical Pneumothorax<br>Simulation Protocol | | | | |
|---|---|---|---|---|
| **Preparation** | | | | |
| Identifies Insertion Site | Locate ipsilateral clavicle | D | ND | DI |
| | Identify 5th intercostal space | D | ND | DI |
| | Transition to the anterior axillary line | D | ND | DI |
| | Mark site avoiding the neurovascular bundle | D | ND | DI |
| Preps patient for Chest Tube insertion | Washes Hands and Wears Gloves | D | ND | DI |
| | Opens kit aseptically | D | ND | DI |
| | Cleans site with ChloraPrep ® in concentric circles | D | ND | DI |
| | Drapes patient with aseptic technique | D | ND | DI |
| Clamp Distal and Proximal End of Chest Tube | Use Kelly clamp to clamp the proximal end of the tube | D | ND | DI |
| | Use Kelly clamp to clamp the distal end of the tube | D | ND | DI |
| **Procedure** | | | | |
| Anesthetize Site | Inject 2% Lidocaine solution into skin incision site | D | ND | DI |
| | Continue to anesthetize deeper subc. tissues and intercostal muscles | D | ND | DI |
| | Insert injection needle into pleural cavity and aspirate air | D | ND | DI |
| Incision | Incise skin 2-4cm parallel to the rib | D | ND | DI |
| Blunt dissection | Bluntly dissect subcu. and intercostal tract above rib | D | ND | DI |
| | Insert finger & assess the tract and rib borders | D | ND | DI |
| | Insert closed Kelly clamp into pleural space | D | ND | DI |
| | Withdraw Kelly clamps with jaws extended | D | ND | DI |
| | Insert finger, sweep 360° and feel for lung or adhesions | D | ND | DI |
| Tube Insertion | Insert tube to proper depth | D | ND | DI |
| Chest Drain Valve | Attach Chest Drain Valve | D | ND | DI |
| Suture Tube in place | Suture tube w/ mattress/interrupted suture | D | ND | DI |
| Apply Dressing | Apply occlusive dressing | D | ND | DI |
| | Dress incision site w/ prepared sterile 4x4 gauze | D | ND | DI |
| | Apply sterile adhesive over incision site | D | ND | DI |

Figure 1. Copy of the Thoracostomy Assessment Tool

After the protocol was compiled and broken down into critical tasks, the procedure was adapted to fit a simulation environment. This process involved removing the following steps: obtain consent from the patient, position patient in a semi-recumbent position, abduct patient's arm over their head, and obtain a chest x-ray. Due to the set up of the mannequin and Chest Tube trainer, the patient mannequin and the mannequin's arm were already in position. Also, verbal cues of consent were removed from the procedure, as they could not be assessed through a visual video recording. The step involving an x-ray was not possible, given simulation set up, and was removed as a follow-up step.

Following the initial adjustments, three expert physicians reviewed the assessment protocol and simulation set up. Recommendations were made about placing the video cameras to capture all of the critical tasks, wording the critical tasks to account for the aseptic environment, and removing unnecessary words. All of the recommendations were incorporated into the final assessment tool. The three expert physicians reviewed the assessment tool a second time and verified the procedure.

*Development of Assessment Scale*

To eliminate subjectivity from the assessment tool, an evaluation scale was adopted to measure competency of each critical step. The scale included three evaluations: Done (D), Not Done (ND), and Done Incorrectly (DI). These evaluations distinguished three levels of competency without requiring the evaluator to judge how incorrect a task was completed. For consistency, each evaluator was asked to mark tasks completed out of order as Done Incorrectly.

*Simulation Set Up*

An Emergency Care Simulator Mannequin was placed on a gurney in a semi-recumbent position to represent the patient. The right arm of the Mannequin was secured to the right corner of the gurney to allow a Wearable Chest Tube Trainer by SynDaver™ Labs to be strapped onto the right side of Mannequin. The procedural area included a sink, gloves, a medical tray, a chest tube, and an aseptic chest tube kit.

Three B-Line Medical cameras were strategically placed to directly capture each critical task. One camera was placed to view the sink and medical tray from behind the procedure. The other two cameras captured an overview of the chest tube trainer from either side of the gurney. The cameras were connected using the web-based SimCapture system to simultaneously record videos of the procedure. The videos could simultaneously be reviewed after the procedure on a single screen through the SimCapture program.

*Experimental Design*

In order to test the objectivity of the assessment tool and compare self, peer, and expert evaluations, 16 inexperienced medical students were recruited in pairs from the University of Arizona College of Medicine. Inexperienced students were chosen to eliminate bias associated with previous thoracostomy competency.

Each student was exposed to the thoracostomy procedure for the first time by watching the New England Journal of Medicine Chest-Tube Insertion video in clinical medicine.[7] Students were able to review the assessment tool while watching the video. After the completion of the video, each student pair was shown a slideshow of the instruments they would have available within the procedural area and in the aseptic chest tube kits. At this time, students were encouraged to ask questions about the instruments, the assessment tool, and the procedure. Following the completion of the questions, the students returned their copy of assessment tool.

Individually, students were brought to the procedural area. The B-Line cameras were turned on, and the students were given permission to begin the thoracostomy procedure. Only simulation- and instrument-specific questions were answered during the procedure. Following the procedure, each student was brought to an isolated location to complete a self and peer assessment using the Thoracostomy Assessment Tool and video recordings of each procedure. The self and peer assessment results for each student were recorded.

An expert evaluator was recruited from Banner University Medical Center to review and assess each procedure. The expert assessment results were recorded as a control for self and peer comparison.

*Statistical Analysis*

The assessment results were analyzed in two ways to measure the accuracy of self and peer assessments and to measure the objectivity of the assessment tool.

In order to measure the accuracy of self and peer assessment in comparison to expert assessments, a binary scale was used. If a variation between evaluation responses existed, the critical task was assigned a value of 1. If both the self-evaluator and expert evaluator had the same evaluation response, the critical task was assigned a value of 0. The same values were assigned for the variations between peer and expert assessments. The binary results for self and peer evaluation variations were pooled into two columns. Two one-way t-tests were completed to determine the statistical significance of the variations.

The objectivity of each critical task in the assessment tool was measured by using the same binary scale from the previous analysis. The self and peer evaluation variations of each procedure were compiled for each of the 25 critical tasks. The data sets for each critical task were analyzed through 25 one-way t-tests.

## Results

*Accuracy of Self and Peer Evaluations*

Two t-tests were used to analyze the significance of variation of self-assessments from expert assessments and peer assessments from expert assessments at a 95% confidence level. The first t-test showed that the variation between self and expert evaluations was significant ($p = <0.00001$). The variation between peer and expert evaluations also produced significant variation ($p = <0.00001$). The null hypothesis that no significant difference exists between self and expert evaluations and peer and expert evaluations was rejected by the self and peer t-tests. The results of the t-tests are displayed in Table 1.

| Evaluation Type | N | Mean | Standard Error | T Statistic | T Critical | DF | P Value | Decision |
|---|---|---|---|---|---|---|---|---|
| Self | 393 | 0.1908 | 0.0198 | 9.6152 | 1.649 | 392 | <0.00001 | Significant |
| Peer | 368 | 0.2255 | 0.0218 | 10.3383 | 1.649 | 367 | <0.00001 | Significant |

Table 1. t-test results for the variation between self, peer, and expert evaluations

*Objectivity of the Assessment Tool*

| Critical Task | N | Mean | Standard Error | T Statistic | T Critical | DF | P Value | Decision |
|---|---|---|---|---|---|---|---|---|
| 1 | 32 | 0.1250 | 0.0594 | 2.1044 | 1.696 | 31 | 0.0218 | Significant |
| 2 | 32 | 0.1563 | 0.0652 | 2.3960 | 1.696 | 31 | 0.0114 | Significant |
| 3 | 32 | 0.2500 | 0.0778 | 3.2146 | 1.696 | 31 | 0.0015 | Significant |
| 4 | 32 | 0.3125 | 0.0832 | 3.7538 | 1.696 | 31 | 0.0004 | Significant |
| 5 | 30 | 0.2667 | 0.0821 | 3.2474 | 1.699 | 29 | 0.0015 | Significant |
| 6 | 28 | 0.4286 | 0.0952 | 4.5000 | 1.703 | 27 | <0.0001 | Significant |
| 7 | 32 | 0.0625 | 0.0435 | 1.4376 | 1.696 | 31 | 0.0803 | Insignificant |
| 8 | 32 | 0.0625 | 0.0435 | 1.4376 | 1.696 | 31 | 0.0803 | Insignificant |
| 9 | 28 | 0.1786 | 0.0737 | 2.4227 | 1.703 | 27 | 0.0112 | Significant |
| 10 | 30 | 0.1333 | 0.0631 | 2.1122 | 1.699 | 29 | 0.0217 | Significant |
| 11 | 32 | 0.0313 | 0.0313 | 1.0000 | 1.696 | 31 | 0.1625 | Insignificant |
| 12 | 32 | 0.1250 | 0.0594 | 2.1044 | 1.696 | 31 | 0.0218 | Significant |
| 13 | 32 | 0.1875 | 0.0701 | 2.6747 | 1.696 | 31 | 0.0059 | Significant |
| 14 | 32 | 0.0313 | 0.0313 | 1.0000 | 1.696 | 31 | 0.1625 | Insignificant |
| 15 | 32 | 0.0938 | 0.0524 | 1.7908 | 1.696 | 31 | 0.0415 | Significant |
| 16 | 32 | 0.2500 | 0.0778 | 3.2146 | 1.696 | 31 | 0.0015 | Significant |
| 17 | 32 | 0.3125 | 0.0832 | 3.7538 | 1.696 | 31 | 0.0004 | Significant |
| 18 | 30 | 0.2667 | 0.0821 | 3.2474 | 1.699 | 29 | 0.0015 | Significant |
| 19 | 32 | 0.4063 | 0.0882 | 4.6055 | 1.696 | 31 | <0.00001 | Significant |
| 20 | 32 | 0.5313 | 0.0896 | 5.9273 | 1.696 | 31 | <0.00001 | Significant |
| 21 | 32 | 0.0938 | 0.0524 | 1.7908 | 1.696 | 31 | 0.0415 | Significant |
| 22 | 32 | 0.5313 | 0.0896 | 5.9273 | 1.696 | 31 | <0.00001 | Significant |
| 23 | 32 | 0.1250 | 0.0594 | 2.1044 | 1.696 | 31 | 0.0218 | Significant |
| 24 | 32 | 0.0313 | 0.0313 | 1.0000 | 1.696 | 31 | 0.1625 | Insignificant |
| 25 | 32 | 0.0625 | 0.0435 | 1.4376 | 1.696 | 31 | 0.0803 | Insignificant |

Table 2. t-test results comparing the variations of peer and self assessments from expert assessments at each critical task

As shown in Table 2, 25 t-tests were used to determine the significance of the variation of the self and peer assessments from expert assessments for each critical task. The analysis was significant across 19 of the 25 critical tasks at a level of 95% significance.

## Discussion

The significant variation found between self and peer evaluations from expert evaluations does not support the use of self and peer assessments in a simulation environment. Although both self and peer t-tests showed significant variations ($p = <0.00001$), fewer variations were found within the self t-test. If assessor subjectivity were reduced throughout the assessment tool, self assessments may be an effective technique to evaluate procedural competency within a simulation environment. Through the use of self assessments, students would have the ability to get accurate feedback without the need for an expert evaluator. This technique for simulation assessment could provide a solution for medical educators limited by time availability. Due to the high levels of variation present at each critical task, self and peer assessments do not produce accurate feedback about student competency and cannot be effectively implemented in simulation assessment using the current Thoracostomy Assessment Tool.

The Thoracostomy Assessment Tool developed in the study was shown to be subjective with significant variation throughout 19 of the 25 critical tasks. Ideally, an objective assessment tool and simulation set up would have no variation among assessors. To help refine the assessment tool to a greater degree of objectivity, the critical tasks with the most variation can be analyzed for improvement. Areas such as camera views and critical task wording can be adjusted to decrease variation. By repositioning the video cameras or adding additional video angles, assessors will have more clarity on the level of completion of critical tasks, such as critical task 6: opens kit aseptically (p = <0.00001). Adding an additional camera above the left side of the mannequin may provide visual clarity to evaluators. Revising the wording of critical tasks may also be a solution to remove subjectivity. Including the definitions of specific phrases, such as "opens kit aseptically" in critical task 6 (p = <0.00001), "proper depth" in critical task 20 (p = <0.00001) and "mattress/interrupted suture" in critical task 22 (p = <0.00001), may clarify the specific actions assessors are being asked to evaluate.

Limitations in simulation assessment emerged throughout the study in relation to clinical preparedness. When the thoracostomy procedure was originally compiled, procedural steps were removed to adjust for the simulation set up. Critical steps in a clinical setting, such as obtaining consent and correctly positioning the patient, were removed from the simulation assessment. By teaching students procedures using adapted assessment tools, they will be unprepared to transfer their procedural competencies directly to a clinical setting without additional training. The inclusion of clinical specific steps on an informational page would help students overcome this educational barrier. Additionally, patient care requires effective communication and interaction.[12] These skills are not neccessary to complete a procedure within a simulation environment but are considerable factors in determining procedural competence.[12] This simulation limitation could have been avoided by using the SynDaver™ Labs Wearable Chest Tube Trainer on a volunteer, instead of a mannequin. Although simulation education cannot fully replicate clinical settings, trainers and assessment tools are improving to better prepare students for clinical environments.[2,12]

When incorporating standardized simulation assessment into an educational program, regular assessment revisions and the inclusion of timed critical tasks should be considered. Assessment tools must be regularly updated to account for technique improvements and new instruments.[6] Revisions could be made annually, or as needed, and updated on a public assessment database.[6] Similarly, the inclusion of timed critical tasks within the tool would allow for a more comprehensive competency assessment.[2] The timed section would assess the completion of specific critical tasks within a certain time span using the evaluation scale of Done and Not Done. Timed sections could be presented as an optional section of the assessment tool, based on learning objectives.

The standardization of simulation assessment is the next step to effectively incorporating simulation education into medical curricula. The method used in this study was unsuccessful at creating a fully objective assessment tool, however, the Thoracostomy Assessment Tool provides a foundation for improvements. If adjustments can be made to reduce subjectivity within the assessment tool, it can be tested and applied to other procedures within the medical field. Through coordinated efforts across simulation research, the development of standardized, objective assessment tools for all procedures in a simulation environment is within reach.

**Endnotes**

[1]Issenberg, S. B. "Simulation Technology for Health Care Professional Skills Training and Assessment." *JAMA: The Journal of the American Medical Association* 282.9 (1999): 861-66. Web.

[2]Scalese, Ross J., Vivian T. Obeso, and S. Barry Issenberg. "Simulation Technology for Skills Training and Competency Assessment in Medical Education." *Journal of General Internal Medicine*. Springer-Verlag, 19 Dec. 2007. Web.

[3]Holmboe, Eric, Mary Ann Rizzolo, Ajit K. Sachdeva, Morton Rosenberg, and Amitai Ziv. "Simulation-Based Assessment and the Regulation of Healthcare Professionals." *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare* 6 (2011): S58-62. Web.

[4]Kim, John, David Neilipovitz, Pierre Cardinal, and Michelle Chiu. "A Comparison of Global Rating Scale and Checklist Scores in the Validation of an Evaluation Tool to Assess Performance in the Resuscitation of Critically Ill Patients During Simulated Emergencies (Abbreviated as "CRM Simulator Study IB")." *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare* 4.1 (2009): 6-16. Web.

[5]Arora, Sonal, Danilo Miskovic, Louise Hull, Krishna Moorthy, Rajesh Aggarwal, Helgi Johannsson, Sanjay Gautama, Roger Kneebone, and Nick Sevdalis. "Self vs Expert Assessment of Technical and Non-technical Skills in High Fidelity Simulation." *The American Journal of Surgery* 202.4 (2011): 500-06. Web.

[6]Bond, W. F., L. Spillane, and The Cord Core Competencies Simulation Group. "The Use of Simulation for Emergency Medicine Resident Assessment." *Academic*

*Emergency Medicine*. Official Journal of the Society for Academic Emergency Medicine, 8 Jan. 2008. Web.

[7]Dev, Shelly P., M.D., Bartolomeu Nascimiento, Jr., M.D., Carmine Simone, M.D., and Vincent Chien, M.D. "Chest-Tube Insertion — NEJM." *New England Journal of Medicine*. Massachusetts Medical Society, 11 Oct. 2007. Web.

[8]Croskerry, Pat. *Patient Safety in Emergency Medicine*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2009. 311-13.

[9]Kesieme, Emeka B., Andrew Dongo, Ndubueze Ezemba, Eshiobo Irekpita, Nze Jebbin, and Chinenye Kesieme. "Tube Thoracostomy: Complications and Its Management." *Pulmonary Medicine* 2012 (2012): 1-10. Web.

[10]Shlamovitz, Gil, MD. "Tube Thoracostomy." Ed. Ryland Byrd. *Medscape* (2014): n. pag. Web.

[11]Salamonsen, MR, F. Bashirzadeh, AJ Ritchie, HE Ward, and DI Fielding. "A New Instrument to Assess Physician Skill at Chest Tube Insertion: The TUBE-iCOMPT." *Thorax* (2014): n. pag. Web.

[12]Epstein, Ronald, M.D. "Assessment in Medical Education — NEJM." *New England Journal of Medicine*. N.p., 25 Jan. 2007. Web.