Pilot Testing of a Novel Surgical Simulator for Endoscopic Zenker's Diverticulotomy

Nathan D. Wiebracht, MD; John P. Giliberto, MD; Charles Myer IV, MD;
Keith Casper, MD; Kaalan E. Johnson, MD

Objectives/Hypothesis: Restrictions on resident work hours and the increasing purview of otolaryngology reduce the efficacy of the traditional surgical training model. With limited case volumes at many institutions and the unique instrumentation of endoscopic Zenker's diverticulotomy (EZD), simulation may be useful to improve training. In this study, a novel surgical simulator for EZD is developed and validated.

Study Design: Simulation model development.

Methods: An EZD model was designed using an intubation trainer and disposable diverticulum inserts. A novel objective structured assessment of technical skill (OSATS) for EZD was developed. Performance of otolaryngology residents on simulations using the OSATS and time to completion were evaluated during an instructional course. Pre- and postencounter surveys were completed. Inter-rater and intrarater reliability were evaluated via blinded video review of resident performance.

Results: Seventeen residents participated (n = 17). Surveys showed confidence improved two points on a five-point scale ($P < .001$), and 94% agreed that the model would improve resident performance with in vivo EZD. More experienced trainees (postgraduate year [PGY] 3–5, n = 11 vs. PGY 1–2, n = 6) had shorter times to completion ($P < .001$) and higher assessment scores on initial attempts ($P = .006$). Both groups showed significant improvements from initial to final attempts on 30-point scales for global rating by 6.2 ± 4.2 (mean ± standard deviation, $P < .001$). The novel OSATS demonstrated fair live/video reliability (κ = 0.40) and inter-rater reliability (κ = 0.44), and moderate intrarater reliability (κ = 0.60).

Conclusions: Pilot testing of an EZD simulator demonstrated acceptability, content validity, and construct validity. A novel OSATS was developed and evaluated. Further investigation of the impact on operative performance and validation of the OSATS in vivo is needed.

Key Words: Simulation, resident, Zenker's diverticulum, low cost, education, task trainer, evaluation.

Level of Evidence: NA

Laryngoscope, 127:592–596, 2017

INTRODUCTION

The case for surgical simulation training is well documented in the medical literature.[1–7] Simulation, at its core, endeavors to anticipate the future needs of the trainee, and allow these needs to be addressed in a controlled environment. This is becoming increasingly useful, as work hour restrictions create a premium on opportunities to train residents, particularly in the surgical disciplines.[8] Residents are at risk of graduating with technical deficiencies due to a less robust operative experience.[1]

The integration of simulation into residency curricula provides a unique adjunct to surgical training by allowing trainees to become familiar with instruments unique to certain procedures, and creating an open environment for learning without risks to the patient. This educational tool is particularly useful at both developing skills in novice residents to perform common procedures, and for all levels of trainees to prepare for high-acuity, low-frequency events. Otolaryngology simulators already exist for myringotomy, tonsillectomy, sinus surgery, airway endoscopy, intubation, and control of epistaxis, with wide variation in cost and fidelity.[9–14]

Endoscopic Zenker's diverticulotomy (EZD) is one example of an infrequently performed procedure in many training programs. Improper use of the Weerda laryngoscope, or the stapler, can lead to perforation and result in significant morbidity and even mortality. A surgical simulator could help to provide an arena for mitigating risk to the patient while allowing the trainee to gain competence with the procedure.

A simulator for EZD was previously developed by Richtsmeier.[15] The model benefitted from being low cost. It also allowed the trainee to become familiar with
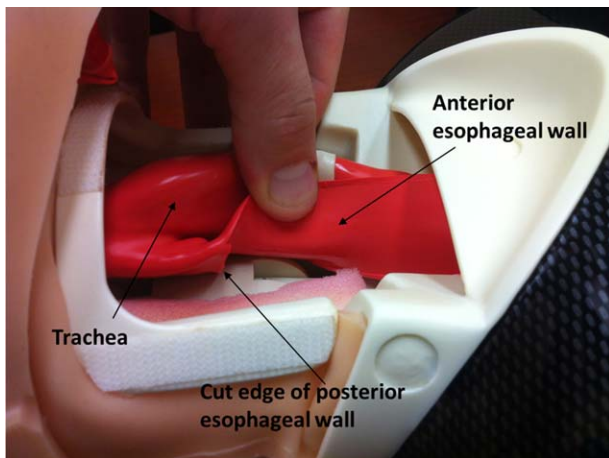
Fig. 1. The posterior half of the blind esophagus on the AirSim intubation manikin is removed to allow accommodation of the prosthetic insert (superior is the left side of figure). [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

instrumentation outside of the operating room. However, because the model was assembled from polyvinyl chloride pipe and plywood, it had some limitations in its face and content validity.

The objectives of this study were twofold, first, to develop a novel simulator for EZD that is low cost and

Fig. 2. The prosthetic inserts. Shown is the glove's fourth finger (top of each image) as the esophagus, and the fifth finger (bottom of each image) as the diverticulum. (A) Glove prosthetic, external surface. (B) Glove prosthetic, "mucosal" surface after dye application. (C) The silicon prosthetic, external surface. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]
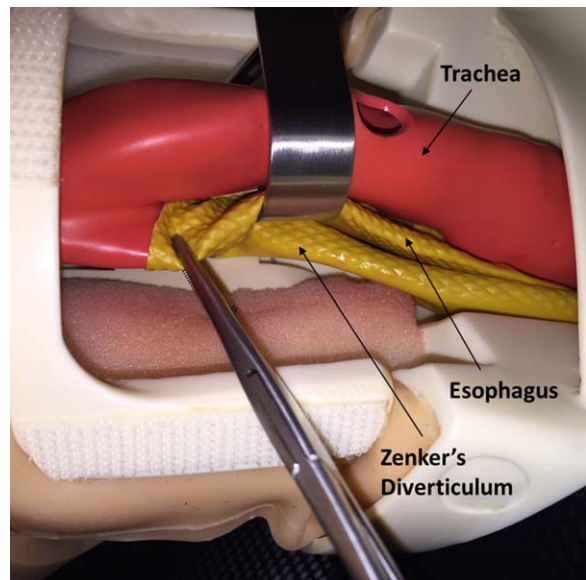

Fig. 3. The prosthetic shown in position on the AirSim intubation manikin. Hemostats are applied at the level of the cricopharyngeus to hold the prosthesis in place during the simulation (retractor is used only for this image to reveal the Zenker's). [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

readily implemented in surgical training programs, and second, to create and validate an objective structured assessment of technical skills (OSATS) as an evaluative tool for trainees performing this procedure.[16–18]

MATERIALS AND METHODS

Model Development

The EZD model used an AirSim intubation manikin (TruCorp Ltd., Belfast, Northern Ireland). The posterior half of the esophagus distal to the piriform recesses was resected (Fig. 1). This allowed placement of a disposable esophagus model containing the Zenker's diverticulum. Adapted from the glove model used by Richstmeier,[15] reusable latex gloves were modified (medium size Caring Hands; Big Time Products, Rome, GA) to create the disposable diverticular model. The fourth finger serves as an extension for the esophagus, and the fifth finger functions as the Zenker's diverticulum. The fourth and fifth fingers, along with 5 cm of the hand portion of the glove, were excised (Fig. 2). The remaining portion of the glove was discarded. The gloves were inverted, and immersed for 3 hours in a solution of 120 mL of Rit scarlet dye (Rit Dye Co., Stamford, CT) mixed with 3.8 L of water. After drying, the gloves were inverted back to their natural position, and a double row of office staples was placed extending superiorly from medial base of the fourth finger. The distal 1 cm of the fourth finger was removed to represent the open distal esophagus. To fit the diverticulum model onto the TruCorp simulator, a hemostat was inserted into the distal end of the fourth finger. The hemostat was used to grasp the TruCorp esophagus and provide traction, while the glove was moved proximally on the native esophagus to the point where the fifth finger was in the position of a Zenker's diverticulum. Two hemostats were used to clamp the glove to the trainer at the position of the cricopharyngeus muscles (Fig. 3).

An alternative diverticulum model was developed in conjunction with a biomedical engineering company (3D Med, Franklin, OH; www.3-dmed.com) in an effort to produce an insert
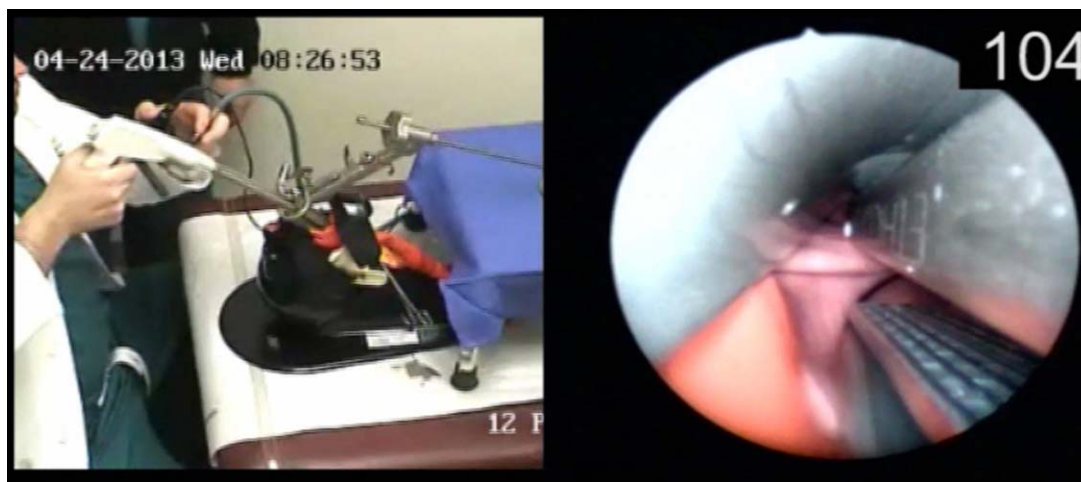
Fig. 4. Screen capture of synchronized video with overhead footage (left) and endoscopic footage (right) for use in remote review. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

with greater face validity and ease of use. Greater anatomic detail was incorporated through an iterative refinement process with the authors to develop a silicone insert model, which was able to be inserted onto the modified esophagus with a tighter fit not requiring hemostats to secure it; this model was available during the project as an alternative to the glove model.

OSATS Development

A novel objective assessment tool was developed entitled the Endoscopic Zenker's Diverticulotomy Objective Evaluation and Skills Inventory Tool (EZ-DOES IT). A task-specific portion was developed by dividing the operative procedure into the specific tasks deemed to be necessary for the completion of a safe and efficient surgery, based upon previously developed anchors.[18] A global-operative performance portion of the scale was developed based on the American College of Surgeons Global rating scale. Task-specific performance and global-operative performance each had six specific aspects that were rated on a five-point Likert scale, for a total score of 30 in each of the two categories. Prior to data collection, all observing faculty underwent a training session, including review of a videotaped simulation using the developed model.

Data Collection

The study was deemed exempt from review by the institutional review board of the University of Cincinnati Medical Center (UCMC). Volunteers were recruited from residents of the UCMC Department of Otolaryngology. All subjects completed a pre- and poststudy self-assessment survey.

All simulations occurred in the Simulation Center at the University of Cincinnati College of Medicine. Stations were equipped with the requisite instruments for the procedure including a 4 mm × 30 mm Hopkins rod telescope (Karl Storz, Tuttlingen, Germany), Weerda laryngoscope (Karl Storz), suspension arm, and an ENDOPATH ETS articulating linear cutting device (ATW35; Ethicon Endo-Surgery, Cincinnati, OH). Each trainee was allowed two rated attempts with a period of directed practice between these attempts. Directed practice was supervised by attending physicians. Trainees were shown proper technique by the attending physician, and each coached on specific ways to improve their individual operative skills in a 1-hour session. Using the EZ-DOES IT OSATS, each attempt was rated under direct observation by an observing faculty member. Each of the two directly observed attempts were video recorded, capturing both a room view and the endoscopic view of the model. Each participant's combined video was synchronized and deidentified, leaving only the participants' hands visible when possible (Fig. 4). Study data were collected and managed using REDCap electronic data capture tools.[19] Combined videos were compiled and evaluated in a randomized order by two faculty raters at two time points, separated by 4 months.

Statistical Analysis

Residents (postgraduate year (PGY) 1 to PGY 5) were dichotomized comparing PGY 1–2 versus PGY 3–5. Performance was evaluated based on time to completion in seconds, and EZ-DOES IT total scores and subscales for task-specific performance and global-operative performance. Improvement was based on changes in time to completion, EZ-DOES IT scores, and changes in resident confidence to perform the procedure as determined by survey responses. Independent t tests were used to compare time to completion and EZ-DOES IT scores by experience level at the initial attempt. Paired t tests were used to compare time to completion and EZ-DOES IT scores between attempts. The Wilcoxon signed rank test was used to assess confidence level by attempt. Significance was assessed at $\alpha = .05$. Cohen's κ was used to assess intrarater reliability, inter-rater reliability, and video-review versus direct-observation reliability of the EZ-DOES IT. All Likert values were assigned a value of 1 to 5, and NA (not able to rate) was assigned a value of 6 for linear κ calculations. SAS version 9.3 (SAS Institute, Cary, NC) was used to conduct the analysis.

RESULTS

Model Validation

A total of 17 residents participated in the simulation activity (PGY 3–5, n = 11 and PGY 1–2, n = 6). There was a 100% response rate for the survey. Most residents agree or strongly agree (94%) that future residents would be better prepared for live EZD after using the simulator. One resident answered "neither agree nor disagree." Mean confidence to perform the procedure significantly improved two Likert points from "basic comfort with steps" to "confident with all steps" ($P < .001$).

TABLE I.
EZ-DOES IT Scoring of Live Attempts.

	First Attempt			Second Attempt		
	Total, n = 17	PGY 1–2, n = 6	PGY 3–5, n = 11	Total, n = 17	PGY 1–2, n = 6	PGY 3–5, n = 11
Time, s	246.3	293	220.8	191.4*	235.2*	167.5*
EZ-DOES IT total, maximum 60	35.47	25.67	40.82	47*	37.67*	52.09*
Task-specific, maximum 30	17.88	13.67	20.18	23.24*	18.83*	25.64*
Global-operative, maximum 30	17.59	12	20.64	23.76*	18.83*	26.45*

*Significant changes, $P < .001$.
PGY = postgraduate year.

Time to completion decreased from the first to the second attempt (Table I). Overall scores significantly improved from the first to the second attempt (mean increase 11.5 ± 9, $P < .001$). Task-specific and global-performance scores also significantly improved.

Trainees with more experience spent less time completing the task compared to those with less experience. Although trainee completion times decreased after the second attempt overall, the change in time did not significantly differ with increasing experience (PGY 3–5 vs. PGY 1–2 mean decrease in time: 53.4 ± 32.7 vs. 57.8 ± 19.9 seconds, $P = .88$).

Performance differed by experience level. Trainees with more experience had higher mean scores compared to trainees with less experience, after the first and second attempts. Although trainee scores improved after the second attempt overall, the change in score did not significantly differ with increasing experience (mean increase PGY 3–5: 12 ± 10.5 vs. PGY 1–2: 11.3 ± 8.6, $P = .88$).

EZ DOES IT Validation

Twelve of the 34 videos only had the overhead room footage because the endoscopic view was not recorded. These 12 had a higher rate of NA. In spite of this, intra-rater reliability based on linear-weighted κ showed substantial agreement for faculty rater 1, and moderate agreement for faculty rater 2 (Table II). Inter-rater reliability of the first and second trainee task attempts showed moderate agreement, whereas comparison of live versus video evaluations of residents showed fair agreement (Table III).

DISCUSSION

The novel surgical simulator for EZD presented and validated in this article appears to demonstrate accepta-

ble content and construct validity. This high-fidelity model builds on a prior low-fidelity model initially developed by Richtsmeier, while keeping the cost of simulation repetitions low. Survey results indicate that 94% of residents believed it would better prepare them for in vivo EZD. This increase in confidence was associated with a statistically significant improvement in objective measures, including time to completion, and both task-specific and global objective ratings of resident performance. Taken together, these results demonstrate that this model could provide a low-risk complement in the surgical training of otolaryngology residents.

Construct validity of the model was supported by the difference between senior versus junior residents, with senior residents significantly outperforming juniors. It is interesting to note that in spite of their higher initial skill level, the seniors improved almost the same magnitude as the juniors. This reinforces the notion that simulation benefits the spectrum of trainees and need not be limited to novices.

The simulator developed for the present study provides significant advances in fidelity, while minimizing the cost of simulation repetition. After an initial investment in the AirSim Intubation Manikin, the prosthetic attachments are less than $2 (USD) for the glove model, and approximately $20 (USD) for the silicone model. Anecdotally the silicone insert model provided improved visual realism and ease of use, but with the large volume of models needed for this project, the low cost and availability of the glove model resulted in that model being used exclusively for data collection. Additionally, the simulator has acceptable ease of use and requires minimal time per attempt, usually less than 5 minutes. Thus, if a program already owns the AirSim Intubation Manikin, a trainee can easily complete multiple iterations relatively easily with low cost.

In addition to creating a novel simulator, the authors sought to establish a rating tool, the EZ-DOES IT. By having faculty raters judge resident performance

TABLE II.
Inter-rater Reliability.

	% Concordance Between Ratings	Cohen's κ	P Value
Video review, FR 1 vs. FR 2	43	0.44	<.00001
Video vs. direct observation	43	0.40	<.00001

FR1 = faculty rater 1, FR2 = faculty rater 2.

TABLE III.
Intra-rater Reliability.

	% Concordance Between Ratings	Cohen's κ	P Value
Faculty rater 1	57.1	0.62	<.00001
Faculty rater 2	54.9	0.59	<.00001

live, followed 4 months later by viewing deidentified synchronized overhead and endoscopic video recorded on the initial day, intrarater reliability was established. Using the video review and EZ-DOES IT, it is quite possible that a trainee could receive remote coaching from a faculty member. This may help to relieve some of the inevitable scheduling obstacles between trainees and trainers. Inter-rater reliability was demonstrated through comparing the two faculty member's ratings. Further refinement of the Likert scale descriptors may aid in improving inter-rater reliability.

This study has several limitations inherent in simulation studies. It was performed in an educational course setting, which produced logistical constraints on available time and participant composition. Malfunctioning video recording equipment resulted in a lower rate of successful recording of the room view for some of the trainee attempts, resulting in a larger number of NA data from these rooms. In spite of this lack of data, however, all of the OSATS score evaluations showed some degree of significant reliability.

The present model developed by the authors may help to accelerate the learning curve such that a novice approaching an in vivo diverticulotomy procedure begins with a greater level of comfort and skill with the procedure than could be gained merely through reading or observing. This then could maximize the opportunity to master technique and increase safety for the patient. Further testing will be needed to demonstrate predictive validity of this model to improve operative performance in vivo. With the preliminarily validated EZ-DOES IT OSATS of the present study, evaluation of content validity will be facilitated in future studies. Higher numbers of participants will be needed to determine which components of the task-specific or global-operative performance benefit the most.

CONCLUSION

A novel low-cost task trainer to simulate EZD was developed, and pilot data demonstrated feasibility, content, and construct validity. Otolaryngology residents showed improvement in task-specific and global-operative performance, and reported increased confidence after use of the model. This model has potential as an adjunct to preoperative surgical training for EZD. Further investigation is needed to determine predictive validity for transferring skills to the operating room and further development of the EZ-DOES IT OSATS.

Acknowledgments

BIBLIOGRAPHY

1. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med* 2006;355:2664–2669.
2. Grantcharov TP, Reznick RK. Teaching procedural skills. *Br Med J* 2008; 336:1129–1131.
3. Schavieren MV. Development of expertise in surgical training. *J Surg* 2010;67:37–43.
4. Sadideen H, Kneebone R. Practical skills teaching in contemporary surgical education: how can educational theory be applied to promote effective learning. *Am J Surg* 2012;204:396–401.
5. Sadideen H, Alvand A, Saadeddin M, Kneebone R. Surgical experts: born or made? *Int J Surg* 2013;11(9):773–778.
6. Cowan JA, Dimick JB, Thompson BG, Stanley JC, Upchurch GR. Surgeon volume as an indicator of outcomes after carotid endarterectomy: an effect independent of specialty practice and hospital volume. *J Am Coll Surg* 2002;195:814–821.
7. Scott DJ, Cendan JC, Pugh CM, Minter RM, Dunnington GL, Kozar RA. The changing face of surgical education: simulation as the new paradigm. *J Surg Res* 2008;147:189–193.
8. Whang EE, Mello MM, Ashley SW, Zinner MJ. Implementing resident work hour limitations. *Ann Surg* 2003;237:449–455.
9. Malekzadeh S, Hanna G, Wilson B, Pehlivanova M, Milmoe G. A model for training and evaluation of myringotomy and tube placement skills. *Laryngoscope* 2011;121:1410–1415.
10. Volsky PG, Hughley BB, Peirce SM, Kesser BW. Construct validity of a simulator for myringotomy with ventilation tube insertion. *Otolaryngol Head Neck Surg* 2009;141:603–608.
11. Fried MP, Sadoughi B, Gibber MJ, et al. From virtual reality to the operating room; the endoscopic sinus surgery simulator experiment. *Otolaryngol Head Neck Surg* 2010;142:202–207.
12. Malekzadeh S, Pfisterer MJ, Wilson B, Na H, Steehler MK. A novel low-cost sinus surgery task trainer. *Otolaryngol Head Neck Surg* 2011;145: 530–533.
13. Deutsch ES, Christenson T, Curry J, et al. Multimodality education for airway endoscopy skill development. *Ann Otol Rhinol Laryngol* 2009; 118:81–86.
14. Malekzadeh S, Malloy KM, Chu EE, et al. ORL emergencies boot camp: using simulation to onboard residents. *Laryngoscope* 2011;121:2114–2121.
15. Richtsmeier WJ. Simulated Zenker's endoscopic staple-assisted esophago-diverticulostomy (ESED) surgery. *Laryngoscope* 2002;112:1230–1234.
16. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;84:273–278.
17. Ahmed K, Miskovic D, Darzi A, Athanasiou T, Hanna GB. Observational tools for assessment of procedural skills: a systematic review. *Am J Surg* 2011;202:469–480.
18. Ishman SL, Benke JR, Johnson KE, et al. 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. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–381.