



Published in final edited form as:

Am J Surg. 2015 January ; 209(1): 132–139. doi:10.1016/j.amjsurg.2014.10.002.

Use of decision-based simulations to assess resident readiness for operative independence

Anne-Lise D. D'Angelo, MD^a, Elaine R. Cohen, MEd^a, Calvin Kwan, BS^a, Shlomi Laufer, PhD^{a,b}, Caprice Greenberg, MD, MPH^a, Jacob Greenberg, MD, EdM^a, Douglas Wiegmann, PhD^c, and Carla M. Pugh, MD, PhD^{a,c}

^aUniversity of Wisconsin School of Medicine and Public Health, Department of Surgery 750 Highland Avenue, Madison, WI 53726-2336; (608) 263-8668, amaag@uwhealth.org

^bUniversity of Wisconsin-Madison, Department of Electrical Engineering and Computer Science 2420 Engineering Hall, 1415 Engineering Drive, Madison, WI 53706; (608)-262-3840; laufer.sh@gmail.com

^cUniversity of Wisconsin-Madison, Department of Industrial and Systems Engineering, 1415 Engineering Drive, Madison, WI 53706, dawiegmann@wisc.edu

Introduction

Recent literature has shown an increasing number of general surgery residents pursuing fellowship training after graduation¹⁻². This trend of seeking additional training may result from residents' lack of confidence in their readiness for independent operative performance³⁻⁷. Nationally, resident readiness has been called into question and is a focus of the ongoing debate regarding the current structure of general surgery training programs³⁻⁴. However, most of the data on graduating residents' ability for independent performance are based on opinion surveys with variable response rates and anecdotal evidence⁵⁻⁸. To better assess resident readiness, evaluations that focus on measuring technical and non-technical skills of graduating residents must be performed.

Additional concerns regarding the structure of general surgery training have come from data relating to operative experiences of surgical residents⁹⁻¹⁰ and increasing failure rates on the Certifying (Oral) Examination of the American Board of Surgery (ABS)¹¹⁻¹². Despite operative case experience requirements by the Accreditation Council for Graduate Medical Education (ACGME), surgery residents have variable operative experiences¹⁰. Of the 121 procedures considered essential by program directors, graduating chief residents in 2005 only performed 18 of them more than 10 times during their general surgery training¹⁰. This data demonstrates that residents are not receiving repeated exposure to a large number of common procedures. It also highlights the need for assessments of surgical performance to determine if residents are reaching competency in essential procedures.

Within the last few years, the ABS has established new requirements for documenting graduating resident skills in the clinical and operative environment. These include passing scores on Fundamentals of Laparoscopic Surgery (FLS) (instituted 2010) and Fundamentals of Endoscopic Surgery (FES) (planned institution in 2017); and documentation of program based-assessments of two clinical and two operative encounters (instituted 2012)¹³. Additionally, residents graduating in 2015 are required to participate in 25 cases as a teaching assistant, where the senior resident instructs a junior resident during a case. These new requirements for performance evaluations and teaching cases during residency training demonstrate the growing concern regarding the need for standardized assessment of and experience in independent operative performance. This is coupled with the fact that the majority of assessment-based research on trainee performance is primarily performed with medical students and junior residents as participants¹⁴.

Current assessment tools for surgical technical skills include task-specific and global rating scales¹⁵; final product analysis¹⁴; and documentation of critical failures¹⁴. The assessment tool Objective Structured Assessment of Technical Skill (OSATS) was originally designed with a task specific checklist, pass/fail judgment and a global rating scale¹⁵. The global rating scale has been advocated by the developers of the OSATS based on data demonstrating higher inter-rater reliability and relationship with experience level as evidence of construct validity¹⁵⁻¹⁷. However there are concerns about the objectivity of the tool in the operating room¹⁸ and data demonstrating lower reliability than procedure based assessments¹⁹. Given the varying data on the use of specific assessment measures and the need for the development of standardized assessment measures of resident readiness for operative independence, we planned to assess graduating chief resident's performance using multiple assessment tools.

In our prior work, we developed error-enabled, decision-based operative simulators²⁰⁻²². These simulators are designed to represent a range of complexity as designated by the Surgical Council on Resident Education (SCORE) curriculum handbook²³. Additionally, the simulation environment allows residents to practice procedures independently and commit errors, which may demonstrate knowledge and skill deficits not identifiable in the operating room. The aim of this study was to assess graduating resident readiness for operative independence on three simulated procedures using multiple assessment measures of performance.

Materials and methods

Setting and Participants

Study participants were graduating chief residents at an academic general surgery training program. Data collection occurred in a single setting approximately one month prior to graduation. This study was reviewed by the University of Wisconsin Health Sciences Institutional Review Board and was approved as a no more than minimal risk study. Written informed consent was obtained from all participants.

Simulated procedures

Chief residents performed three simulated surgical procedures. The procedures were selected to assess both open and laparoscopic skills and evaluate performance on a range of operative complexity as designated by the SCORE curriculum²³. The simulated procedures were: 1) laparoscopic ventral hernia (LVH) repair (laparoscopic essential-common), 2) hand sewn bowel anastomosis (open essential-common) and 3) pancreaticojejunostomy (complex) (Figure 1). The pancreaticojejunostomy was selected for the complex procedure as residents at the institution rotate on a service with multiple pancreatic surgesons.

Participants were given 30 minutes to complete each procedure. The simulators were designed to induce decision-making throughout the procedures. At the bowel anastomosis and pancreaticojejunostomy stations, a faculty member served as an operative assistant. The faculty members had not previously operated with the participants and were instructed to not provide feedback during the procedure. At the LVH station, a researcher with basic laparoscopic skills acted as the operative assistant and did not provide feedback. To avoid sequence effects, the residents did not perform the procedures in the same order. The participants began at either the LVH, bowel anastomosis or pancreaticojejunostomy stations and rotated through the remaining stations in a random order.

Laparoscopic ventral hernia (LVH) repair—This station was designed to represent a patient with a ventral hernia after a midline incision. The performance goal was to successfully repair the LVH. The simulator is a box-style trainer with a plastic base covered by a fabric skin that contains a 10 × 10cm ventral hernia defect^{20–22}. It was designed to look like the abdominal cavity, specifically from the diaphragm to the upper plane of the pelvic cavity. The simulator was fabricated to represent the layers of the abdominal wall including: skin, subcuticular tissue, and peritoneum. Inside, simulated organs including bowel and omentum are layered within the cavity to provide the appearance of gross anatomy. The simulator allows for the use of both laparoscopic and open surgical instruments, except cautery, during each step of the procedure. During this simulated procedure, residents are required to make decisions throughout the procedure including port location, mesh sizing, mesh orientation and mesh fixation.

Bowel anastomosis—This station was designed to represent a trauma patient who sustained a gunshot wound to the abdomen. The goal was to successfully repair an injury to the small bowel. The simulator consisted of porcine intestines arranged in a basin with accompanying artificial blood and retained bullet. The bowel injury was created at the anti-mesenteric border. There were two full thickness injuries in close proximity: one large and one small. The larger injury had jagged edges and was closer to the anti-mesenteric border. The smaller injury was positioned near the mesenteric border of the larger injury. An example of the injury is shown in Figure 2. A complete surgical tray with all the necessary open surgical instruments and suture were provided to complete the repair. During the bowel repair simulation, residents must make operative decisions, including identification of the injuries, choice of repair technique (primary repair vs resection with side-to-side anastomosis), and selection of suture and stitch technique.

Pancreaticojejunostomy—This station was designed to represent a patient with pancreatic cancer along with biliary and pancreatic ductal dilation. The goal was to perform a duct-to-mucosal pancreaticojejunostomy. As such the participants were informed that the resection had already been completed. The simulator consisted of a fabricated pancreas, pancreatic duct, bowel and mesentery arranged within a basin lined with simulated peritoneum^{20,22}. The simulator was constructed from fabric and plastic materials. A complete surgical tray with all the necessary open surgical instruments and suture were provided to complete the repair. During this simulated procedures, residents are required to make operative decisions including planning of anastomosis given relevant anatomy, selection of suture, and alterations in technique in response to simulated tissue characteristics.

Surveys

Prior to performing the simulated procedures, residents completed a survey designed to collect information about demographics, and choice of fellowship specialty or general surgery practice. Operative case logs were obtained to evaluate the number of cases residents had performed during their general surgery training. The pre-procedure survey also utilized a 5 point-Likert scale to assess resident confidence performing each of the three procedures (LVH, bowel anastomosis and pancreaticojejunostomy) independently. After performing each of the simulated procedures, residents completed another survey rating overall task difficulty (5 point-Likert scale), post-procedure confidence (5 point-Likert scale) and the most challenging step in each procedure.

After completing the exit examination, residents also completed a post-assessment survey rating their satisfaction with the simulations (5-point Likert scale) and which year they felt these assessments should take place during their residency.

Performance rating scales

Faculty rated performance using: 1) task-specific checklists, 2) adapted OSATS and 3) final product analysis. Task specific checklists were used for each of the individual stations to evaluate procedure specific technical skills (Appendix). The LVH and pancreaticojejunostomy checklists were developed based on prior cognitive task analysis²² and faculty evaluations²⁰ and the bowel anastomosis checklist was used as previously published¹⁷. The previously published OSATS rating scale was adapted to assess general technical skills on each of the three stations¹⁷. This adapted OSATS scale differed by including assessment of *ability to adapt to individual pathological circumstances* and excluding assessment of *knowledge of specific procedure*. This adaption was made as all of the senior residents had previously performed these procedures and we wanted to focus on surgical planning and ability to adapt to different clinical presentations. Expert faculty surgeons at each station rated resident performance throughout the procedure using the adapted OSATS scale and task specific checklists. Final product analysis of the simulated procedures was performed and faculty members documented errors in technique. Final product analysis of the LVH repair included analysis of hernia coverage, use of anchoring sutures and placement of tacks. For the bowel anastomosis, the final repair was evaluated for technique utilized, orientation of the repair and luminal narrowing. Lastly, the

pancreaticojejunostomy was evaluated for suture selection, injury to the superior mesenteric vein and anterior and posterior suture lines.

Data analysis

We evaluated performance on three simulated operative procedures with rating scales for general technical skills, task specific technical skills and final product analysis. This was done to compare overall performance on different measures and identify deficiencies in operative skills. Pre- and post-procedure confidence ratings were averaged and means were compared with repeated measures analysis of variance (ANOVA) and paired t-tests. Completion rates were determined using the number of residents who finished each procedure. Task-specific and OSATS checklist scores were averaged across stations. To determine which aspects of the OSATS global rating scale residents performed significantly different on, repeated measures ANOVA and follow up t-tests with a Bonferonni correction for multiple pairwise comparisons was performed. Deficiencies were tallied and noted for each procedure. Resident responses on the post-assessment survey were averaged.

Results

Demographics

Six graduating chief residents completed this study. Half of the study participants were male. All of the graduating residents planned to complete fellowships. Fellowship specialties included critical care (N=2), minimally invasive surgery (N=1) and colorectal surgery (N=3). The participants logged total major cases ranging from 870-1162 ($M=974.8$ ($SD=103.1$)). They also logged an average of 15.0 ($SD=2.6$) LVH repairs; 20.7 ($SD=6.1$) open small bowel enterectomies; and 12.5 ($SD=2.6$) total pancreas cases of which on average 6.8 ($SD=2.8$) were pancreaticoduodenectomy (Whipple) procedures. Total teaching assistant cases ranged from 21 to 109 ($M=54.7$, $SD=29.3$).

Pre and post-procedure Confidence

Both pre- and post-procedure, residents were more confident to independently perform the LVH (pre $M=4.0/5.0$, $SD=.63$; post $M=4.0/5.0$; $SD=0.63$) and bowel anastomosis (pre $M=4.3/5.0$; $SD=.52$; post $M=3.8/5.0$, $SD=1.7$) than the pancreaticojejunostomy (pre $M=2.5/5.0$; $SD=.84$; post $M=2.8/5.0$; $SD=1.17$) ($F(1,5)=21.67$, $p<.001$; all pairwise comparisons $p<.01$). There was no significant change in confidence ratings after performing the procedures ($F(1,10)=.056$, $p=.822$).

LVH Repair

Only one of six residents completed the LVH procedure station in the allotted time. Individual LVH task specific checklist scores ranged from 25.0% to 100.0% ($M=61.1\%$; $SD=25.1\%$) (Table 1). Out of 12 independent items on the LVH task specific checklist, the most prevalent incomplete tasks were: *orientating mesh with at least a 3cm overlap on each side* (83.3% incomplete); *placing mesh flat* (83.3% incomplete); and *applying tacks at 1cm intervals to prevent exposure of mesh's adhesiogenic surface* (83.3% incomplete). Mean OSATS score was 4.24/5.0, $SD=.53$. On the OSATS global rating scale, residents scored the lowest for the *use of assistants* and *ability to adapt to individual pathological circumstances*.

Residents scored the highest on respect for tissue (Table 1). Residents most frequently rated *visualizing the hernia defect* (50.0%) as the most challenging procedure step. Final product analysis revealed errors in the placement of tacks and use of an inappropriate number of anchoring sutures.

Bowel anastomosis

All six residents completed the bowel anastomosis procedure station in the allotted time. Individual bowel anastomosis task specific checklist scores ranged from 68.1% to 100.0% ($M=86.1\%$; $SD=13.1\%$) (Table 1). Out of 22 independent items on the bowel anastomosis task specific checklist, the most prevalent incomplete items were *using hemostats to hold stay sutures* (50.0% incomplete); and *placement of inverting sutures* (50.0% incomplete). Mean OSATS score for the bowel anastomosis was 4.08/5.0, $SD=.41$. On the OSATS global rating scale, residents scored the lowest on *respect for tissue* and scored the highest on *knowledge of instruments* (Table 1). Residents most frequently rated *selecting the correct stitch* (50.0%) as the most challenging procedure step. Final product analysis revealed several errors including closure in the longitudinal direction rather than horizontal direction with narrowing of the lumen, suboptimal spacing of suture and missed injury.

Pancreaticojejunostomy

All six residents completed the pancreaticojejunostomy procedure station in the allotted time. Individual pancreaticojejunostomy task specific checklist scores ranged from 68.8% to 93.8% ($M=82.3\%$; $SD=10.8\%$) (Table 1). Out of 16 independent items on the pancreaticojejunostomy task specific checklist, the most prevalent incomplete items were *placing backwall sutures first* (66.7% incomplete); and *use of forceps to handle needle* (50.0% incomplete). Mean OSATS score for the pancreaticojejunostomy was 4.06/5.0, $SD=.51$. On the OSATS global rating scale, residents scored the lowest on *time and motion* and scored the highest on *knowledge of instruments* (Table 1). Residents most frequently rated *planning for backwall sutures* (66.7%) as the most challenging procedure step. Final product analysis revealed errors including failure to place back wall sutures, failure to take pancreatic tissue in bites, puncture of the superior mesenteric vein, and tearing through pancreatic tissue.

Comparative OSATS-analysis

The repeated measures ANOVA on the OSATS global rating scale revealed the main effect for simulation station was non-significant, $F(2,12)=.786$, $p=.478$, indicating that resident scores on the OSATS global rating scale items were not significantly different between the three simulation stations. However, there was a significant main effect for OSATS global rating scale items, $F(2.3, 14.0)=4.29$, $p=.031$ indicating that resident scores on the individual OSATS items were significantly different. Residents received significantly higher scores on *instrument knowledge* ($M=4.78$, $SD=.23$) than *time and motion* ($M=3.94$, $SD=.48$), $t(5)=5.41$, $p=.025$ and *ability to adapt to individual pathological circumstances* ($M=4.06$, $SD=.12$), $t(5)=8.30$, $p=.002$ (Table 1).

Resident evaluation

Likert scale results show a trend towards higher ratings for the bowel anastomosis simulation, followed by LVH then the pancreaticojejunostomy. Residents also felt that they were well prepared to make intra-operative decisions independently (Table 2). Residents almost unanimously agreed that the LVH and bowel anastomosis assessments should take place during their third to fourth year of residency. In addition, almost all residents agreed that the pancreaticojejunostomy assessment should take place during their fifth year.

Discussion

New requirements from ABS for documenting graduating resident skills in the clinical and operative environment have highlighted the need for standardized assessments of resident operative performance. In this study, six graduating chief residents performed three simulated, operative procedures: LVH repair, bowel anastomosis, pancreaticojejunostomy. The aim of this study was to assess graduating resident readiness for operative independence on three simulated procedures using multiple assessment measures of performance.

Only one resident completed the LVH repair in the allotted time, whereas all six participants completed the bowel anastomosis and pancreaticojejunostomy. Residents were most confident in their ability to perform the bowel anastomosis and least confident to perform the pancreaticojejunostomy prior to the procedures. This was consistent with the average number of cases residents logged for each of these procedures prior to participating in this study. Therefore, residents had higher self-confidence in procedures they had performed more frequently. There was little change in overall self-confidence following the simulated procedures; however, post-procedure, residents were most confident in their ability to perform the LVH repair despite only one resident successfully completing this procedure. This indicates that the use of resident confidence as a proxy for readiness for independence may not be the best method for investigating ability for independent operative performance of graduating residents. This is consistent with prior literature revealing that self-assessment alone is not an accurate measure of physician competence²⁴.

Faculty used task-specific checklists, OSATS and final product analysis to rate resident performance. There was variable performance on all three procedures based on task-specific checklist completion (LVH 25.0-100.0%; Bowel anastomosis 68.1% to 100.0%; and pancreaticojejunostomy 68.8% to 93.8%) indicating a range of resident ability. These low performance statistics is in contrast to the relatively high mean OSATS ratings for each of the procedures (LVH 4.23; Bowel anastomosis 4.08; Pancreaticojejunostomy 4.06). Additionally, resident scores on the individual OSATS global rating scale items were not significantly different between the three simulation stations. This indicates that the individual OSATS global rating scale items were not sensitive to differences in performance on the different simulations. Lastly, final product analysis revealed a range of critical errors not captured by the task-specific checklist or OSATS scores. The uniformity in OSATS scores may relate to the tendency for normative-based assessment with this scale, rating performance according to PGY year rather than actual performance²⁵. Additionally, the range of scores on the different measures may result from each assessment tool evaluating a different construct within the domain of independent operative performance. OSATS is

designed to evaluate skills globally, whereas the task specific checklists focus on whether specific components of the procedure are completed. Final product analysis is necessary to ensure that the actual procedure was completed without critical errors. Thus, to best evaluate resident readiness, multiple rating methods need to be used during performance assessment.

On the OSATS rating scale, residents scored significantly higher on *instrument knowledge* than *time and motion* and *ability to adapt to individual pathological circumstances*. *Time and motion* and *ability to adapt* may represent performance domains that are more difficult for residents to master or require additional experience. There was a non-significant trend towards lower OSATS scores for *use of assistants*. The lower scores for *use of assistants* may reflect a lack of experience providing direction in the operating room. Although residents have the opportunity later in their training to serve as a teaching assistant, this experience may be variable as evidenced by the wide range of teaching assistant cases logged (range 21 to 109) by the participants in this study. The ability to direct an operative assistant is an important component of independent performance and a greater effort to provide this experience during residency may be warranted.

The limitations of this study include the small sample size of six participants. This likely limited the power required to find significant differences in performance and correlations between prior experience and performance. Our participants were comprised of the entire graduating class of chief residents at an academic training institution. The consistently small class size of graduating residents at training institutions calls for future work to be pursued at a multi-institutional level to have an adequate sample size. Another limitation of this study was the potential for observer bias²⁶. The use of checklist and OSATS forms with evidence of validity was intended to help mitigate potential observer bias. Lastly, there are limitations in the use of simulation to assess operative performance. The LVH and pancreaticojejunostomy simulators have been previously assessed for evidence of validity²⁰⁻²². The bowel anastomosis assessment was adapted from a previously published checklist with validity evidence¹⁷. Additionally, this simulator was rated highly by participants in terms of usefulness and recommended for assessment during PGYs 3-4. The fact that only one resident completed the LVH procedure in the allotted time could represent difficulty using the simulator. However, our prior work with this particular simulator has demonstrated that senior residents commit fewer errors and have an improved ability to complete the procedure in the allotted time following didactic training²¹. Future work using an error based framework may allow for further characterization of the failure to complete the simulated LVH repair.

This study was performed just prior to graduation, which did not allow for curricular changes to affect the participants. Although feedback was provided to the program coordinator and director, the timing of the assessment underscores the importance of moving competency-based evaluations earlier in residency. The early identification of deficits in knowledge and skill allows for remediation and reassessment throughout the training program. Repeat assessments make tracking of resident progression towards independent operative performance possible.

As surgery faculty and program directors responsible for resident education continue to evaluate the surgical residency curricula, more work is required to assess resident readiness for operative independence. This study is revealing as it used surgical simulators that allow for independent operative performance, direction of operative assistants and intra-operative decision making. Simulation-based assessments of this nature should employ multiple rating scales as task-specific checklists, global rating scale and final product analysis may each assess different constructs within independent surgical performance. Lastly, further development and integration of error-based assessment frameworks may be particularly valuable to identify specific deficits not characterized by checklists or global rating scales.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Funding for this study came from the Department of Defense grant #W81XWH-13-1-0080 entitled, "Psychomotor and Error Enabled Simulations: Modeling Vulnerable Skills in the Pre-Mastery Phase" and the National Institutes of Health grant #1F32EB017084-01 entitled "Automated Performance Assessment System: A New Era in Surgical Skills Assessment".

References

1. Borman KR, Vick LR, Biester TW, et al. Changing demographics of residents choosing fellowships: longterm data from the American Board of Surgery. *J Am Coll Surg*. 2008; 206(5):782–788. [PubMed: 18471695]
2. Ellis MC, Dhungel B, Weerasinghe R, et al. Trends in research time, fellowship training, and practice patterns among general surgery graduates. *J Surg Educ*. 2011; 68(4):309–312. [PubMed: 21708369]
3. Richardson, JD. [Accessed March 6th, 2014] ACS transition to practice program offers residents additional opportunities to hone skills. *The Bulletin of the American College of Surgeons*. <http://bulletin.facs.org/2013/09/acs-transition-to-practice-program-offers-residents-additional-opportunities-to-hone-skills/>. Published 2013.
4. Hoyt, BD. [Accessed March 6th, 2014] Executive Director's report: Looking forward - February 2013. *Bulletin of the American College of Surgeons*. 2013. <http://bulletin.facs.org/2013/02/looking-forward-february-2013/>. Published 2014.
5. Coleman JJ, Esposito TJ, Rozycki GS, et al. Early subspecialization and perceived competence in surgical training: are residents ready? *J Am Coll Surg*. 2013; 216(4):764–771. [PubMed: 23521960]
6. Nakayama DK, Taylor SM. SESC Practice Committee survey: surgical practice in the duty-hour restriction era. *Am Surg*. 2013; 79(7):711–715. [PubMed: 23816005]
7. Bucholz EM, Sue GR, Yeo H, et al. Our trainees' confidence: results from a national survey of 4136 US general surgery residents. *Arch Surg*. 2011; 146(8):907–914. [PubMed: 21844434]
8. Mattar SG, Alseidi AA, Jones DB, et al. General surgery residency inadequately prepares trainees for fellowship: results of a survey of fellowship program directors. *Ann Surg*. 2013; 258(3):440–449. [PubMed: 24022436]
9. Malangoni MA, Biester TW, Jones AT, et al. Operative experience of surgery residents: trends and challenges. *J Surg Educ*. 2013; 70(6):783–788. [PubMed: 24209656]
10. Bell RH Jr, Biester TW, Tabuenca A, et al. Operative experience of residents in US general surgery programs: a gap between expectation and experience. *Ann Surg*. 2009; 249(5):719–724. [PubMed: 19387334]

11. Falcone JL, Hamad GG. The American Board of Surgery Certifying Examination: a retrospective study of the decreasing pass rates and performance for first-time examinees. *J Surg Educ.* 2012; 69(2):231–235. [PubMed: 22365871]
12. Sachdeva AK, Flynn TC, Brigham TP, et al. Interventions to address challenges associated with the transition from residency training to independent surgical practice. *Surgery.* 2014; 155(5):867–882. [PubMed: 24656857]
13. The American Board of Surgery. [Accessed March 28th, 2014] Training Requirements – General Surgery Certification. http://www.absurgery.org/default.jsp?certgsqe_training. Published 2014.
14. Schmitz CC, DaRosa D, Sullivan ME, et al. Development and verification of a taxonomy of assessment metrics for surgical technical skills. *Acad Med.* 2014; 89(1):153–161. [PubMed: 24280844]
15. Martin JA, Regehr R, Reznick H, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 1997; 84:273–278. [PubMed: 9052454]
16. MacRae H, Regehr G, Leadbetter W, et al. A comprehensive examination for senior surgical residents. *Am J Surg.* 2000; 179(3):190–193. [PubMed: 10827317]
17. Reznick R, Regehr G, MacRae H, et al. Testing technical skill via an innovative "bench station" examination. *Am J Surg.* 1997; 173(3):226–230. [PubMed: 9124632]
18. Hiemstra E, Kolkman W, Wolterbeek R, et al. Value of an objective assessment tool in the operating room. *Can J Surg.* 2011; 54(2):116–122. [PubMed: 21251415]
19. Beard JD, Marriott J, Purdie H, et al. Assessing the surgical skills of trainees in the operating theatre: a prospective observational study of the methodology. *Health Technol Assess.* 2011; 15(1):1–162. i-xxi,
20. Pugh CM, DaRosa DA, Santacaterina S, et al. Faculty evaluation of simulation-based modules for assessment of intraoperative decision making. *Surgery.* 2011; 149(4):534–542. [PubMed: 21216420]
21. Pugh C, Plachta S, Auyang E, et al. Outcome measures for surgical simulators: is the focus on technical skills the best approach? *Surgery.* 2010; 147(5):646–654. [PubMed: 20403517]
22. Pugh CM, Santacaterina S, DaRosa DA, Clark RE. Intra-operative decision making: more than meets the eye. *Journal of Biomedical Informatics.* 2011; 44:486–496. [PubMed: 20096376]
23. Surgical Council on Resident Education. [Accessed April 9, 2014] Curriculum Outline for General Surgery Residency 2013–2014. <http://www.absurgery.org/xfer/curriculumoutline2013-14.pdf>. Published 2013.
24. Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. *JAMA.* 2006; 296(9):1094–1102. [PubMed: 16954489]
25. Ghaderi I, Manji F, Park YS, et al. Technical skills assessment toolbox: A review using the unitary framework validity. *Annals of Surgery.* 2014; 0(0):1–12.
26. Angrosino, Michael V. Observer Bias. In: Lewis-Beck, MS.; Bryman, A.; Liao, TF., editors. *Encyclopedia of Social Science Research Methods.* Thousand Oaks, CA: SAGE Publications, Inc.; 2004. p. 758-759.

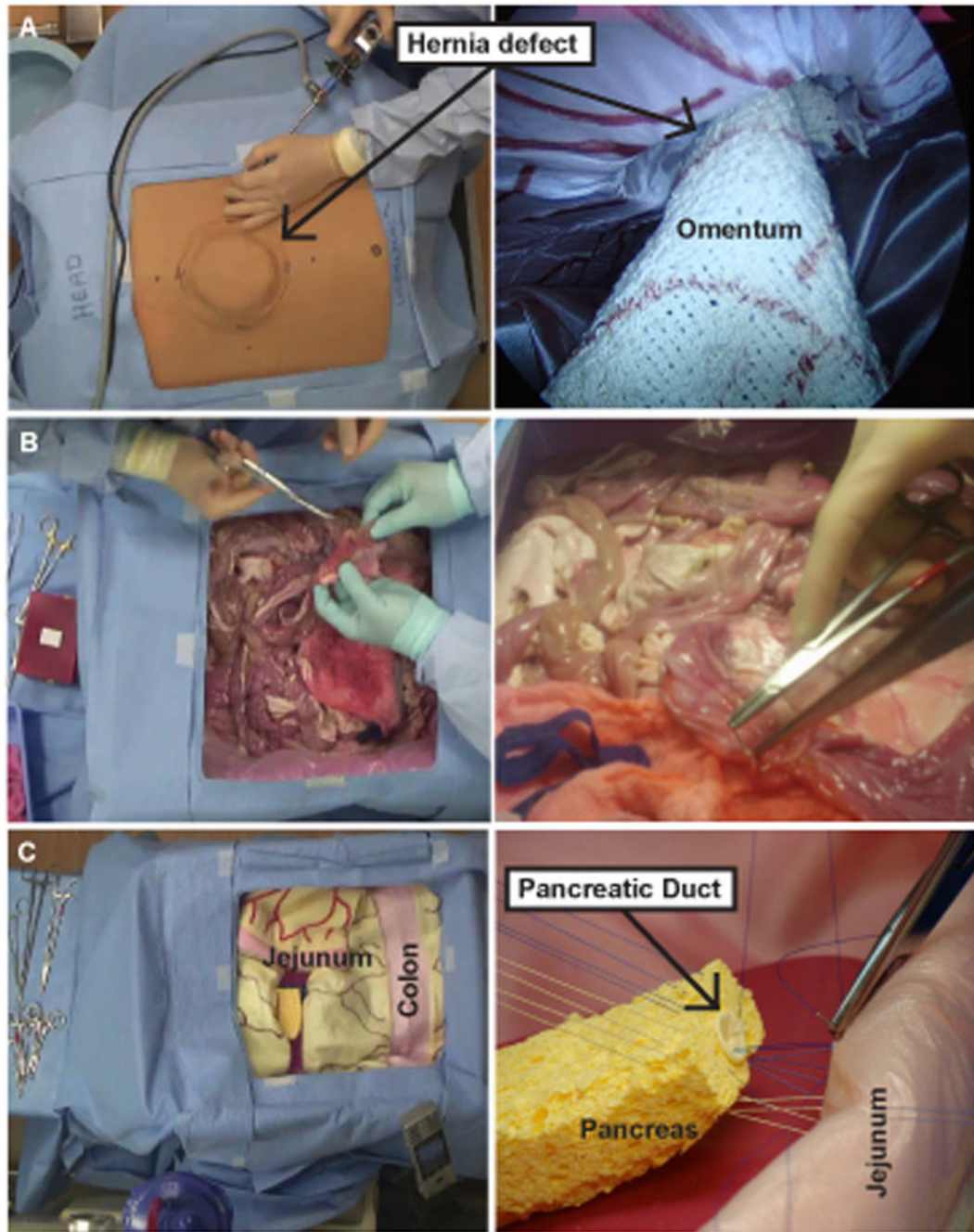


Figure 1. Laparoscopic ventral hernia (A), bowel anastomosis (B), and pancreaticojejunostomy (C) simulators.

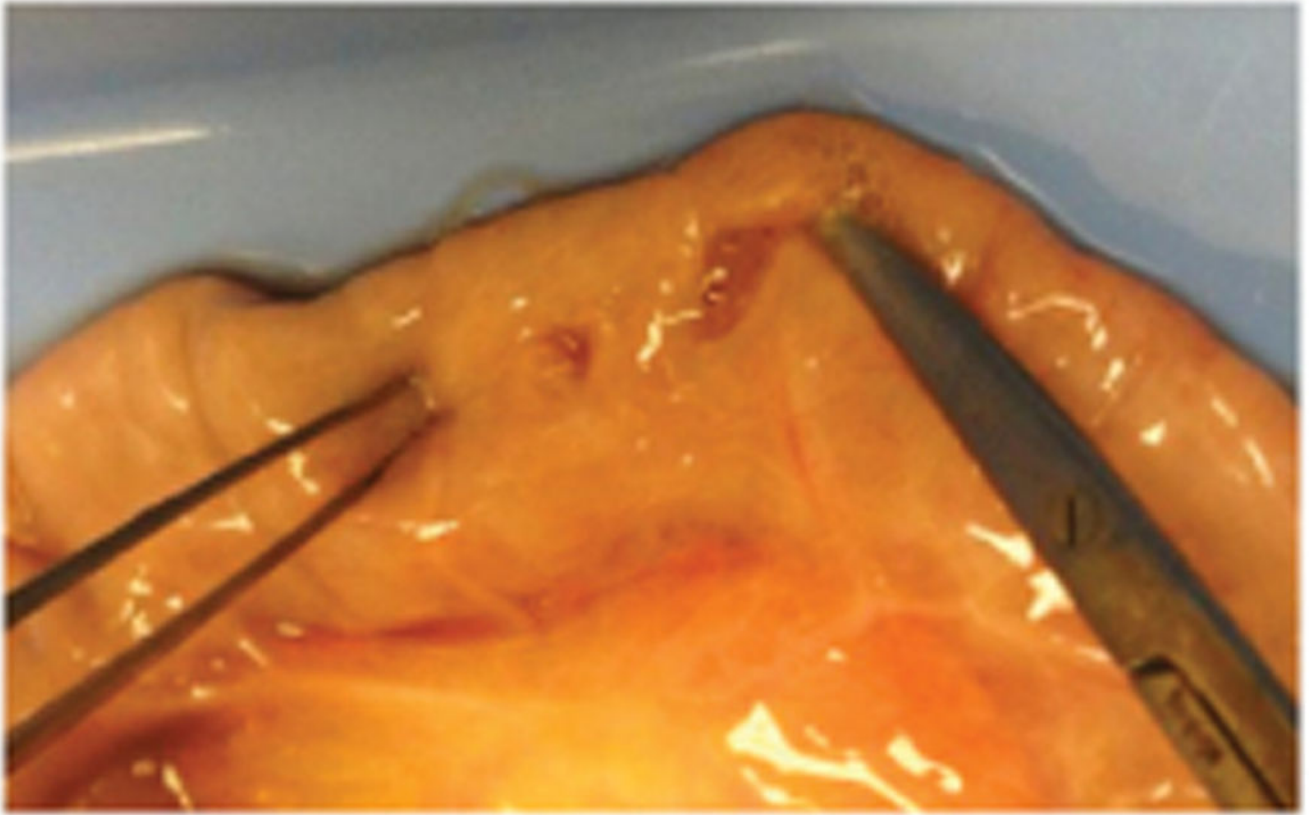


Figure 2.
Example of the created injury for the bowel anastomosis station.

Table 1

Faculty ratings of resident performance for task-specific checklists and OSATS*

Measure	LVH Mean (SD)	Bowel Anastomosis Mean (SD)	Pancreatico- jejunostomy Mean (SD)	Mean Score Across Stations Mean(SD)
Task-specific checklists (%) correct	61.1% (25.1%)	86.1% (13.1%)	82.3% (10.8%)	—
OSATS				
Respect for tissue	4.67 (.82)	3.58 (.80)	4.00 (.63)	4.08 (.84)
Time and motion	4.17 (.41)	3.83 (1.03)	3.83 (.75)	3.94 (.75)**
Knowledge of Instruments	4.83 (.41)	4.83 (.41)	4.67 (.52)	4.78 (.43)**
Use of assistants	4.00 (.89)	3.83 (.41)	3.83 (.98)	3.89 (.76)
Flow of operation and forward planning	4.33 (.82)	4.42 (.66)	4.12 (.98)	4.31 (.79)
Ability to adapt to individual pathological circumstances	4.00 (.63)	4.00 (0)	4.12 (.41)	4.06 (.42)**
Overall performance	3.67 (1.03)	4.08 (.66)	3.75 (.88)	3.83 (.84)

* Responses on a 5-point Likert scale from 1=very poor to 5=clearly superior

** p < .05. Repeated measures ANOVA with follow up t-tests, Bonferonni correction for multiple pairwise comparisons

Table 2

Attitudes* toward exit exam curriculum

Survey Item	Mean (SD)
I received useful educational feedback from the skills assessments	3.33 (1.21)
The LVH simulation model is useful for skills assessment	3.83 (.41)
The bowel anastomosis simulation model is useful for skills assessment	4.17 (.41)
The pancreaticojejunostomy simulation model is useful for skills assessment	3.67 (.82)
Assessments like these should be a required part of the surgery residency curriculum	3.67 (.52)
In general, I feel well prepared to make intra-operative decisions independently	4.10 (.75)

* Responses on a 5-point Likert scale from 1=strong disagree to 5=strongly agree

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript