Using a Simulated Surgical Skills Station to Assess Laceration Management by Surgical and Nonsurgical Residents

MATTHEW V. FARGO, MD, MPH
JOHN A. EDWARDS, MD, MPH, FAAFP
BERNARD J. ROTH, MD, FACP, FACC
MATTHEW W. SHORT, MD, FAAFP

Abstract

Objective To assess laceration management performance among surgical and nonsurgical postgraduate year-1 (PGY-1) residents objectively and to test for interval improvement.

Methods From 2006 to 2008, 106 PGY-1 residents from 10 medical specialties were evaluated with a simulated surgical skills station using pigs' feet before and after internship. Subjects were given 11 minutes to choose the proper suture, prepare and close the wound, and answer laceration management questions. Trainees were classified as surgical (emergency medicine, general surgery, obstetrics and gynecology, orthopedics, and otolaryngology) and nonsurgical (family medicine, internal medicine, neurology, pediatrics, and transitional year). An objective checklist was used to assess performance.

Results A total of 106 PGY-1 residents (age range, 25–44 years; mean, 28.7 years) participated, consisting of 41 surgical (39%) and 65 nonsurgical residents (61%). Surgical group scores improved from 78.4% to 87.7% ($P < .001$). Nonsurgical scores improved from 67.2% to 73.1% ($P < .001$). There was similar improvement between groups (surgical, 9.4%; nonsurgical, 5.9%; $P = .21$). Surgical residents outscored nonsurgical residents before ($P < .001$) and after ($P < .001$) internship.

Conclusion Surgical residents outperformed nonsurgical residents before and after the PGY-1 year with similar score improvements. A simulated surgical skills station can be used to evaluate procedure performance objectively and to test for interval improvement. A simulated surgical skills station may serve as a useful adjunct to apprenticeship in assessing procedure competence.

Matthew V. Fargo, MD, MPH, is Director of the Family Medicine Residency Program at Eisenhower Army Medical Center; **John A. Edwards, MD, MPH, FAAFP**, is Director of the Family Medicine Residency Program at Madigan Healthcare System and Clinical Assistant Professor of Family Medicine at the University of Washington School of Medicine; **Bernard J. Roth, MD, FACP, FACC**, is Pulmonary Disease Subspecialty Education Coordinator at Madigan Healthcare System, Professor of Medicine at the Uniformed Services University of Health Science, and Clinical Professor of Medicine at the University of Washington, Division of Pulmonary/Critical Care Medicine; and **Matthew W. Short, MD, FAAFP**, is Director of the Transitional Year Program and Family Medicine Colonoscopy Fellowship at Madigan Healthcare System, Adjunct Assistant Professor of Family Medicine at the Uniformed Services University of the Health Sciences School of Medicine, and Clinical Assistant Professor of Family Medicine at the University of Washington School of Medicine.

The views expressed are those of the authors and do not reflect the official policy of the Department of the Army, the Department of Defense, or the US Government.

A poster presentation of this work was presented on April 25, 2008, at Western Regional Medical Command Madigan Research Day in Fort Lewis, WA.

Funding: The authors report no external funding source.

Corresponding author: Matthew W. Short, MD, FAAFP, MCHJ-CLE-G, Madigan Healthcare System, 9040 Fitzsimmons Dr, A, Tacoma, WA 98431-1000, 253.968.5017, matthew.w.short@us.army.mil

Received October 20, 2010; revision received December 21, 2010, and March 21, 2011; accepted March 4, 2011.

DOI: <http://dx.doi.org/10.4300/JGME-D-10-00208.1>

Introduction

Skin lacerations are common and result in approximately 7 million to 9 million visits to emergency departments in the United States each year.¹ A baseline knowledge of wound closure and management is expected of physicians in many medical specialties for laceration repair, minor skin procedures, and major surgery. Graduate medical education programs are faced with the daunting task of ensuring resident procedural competence through apprenticeship and evaluation.

Surveys have shown that residents feel less experienced in performing basic procedures,² and program directors have called for more standardized methods for teaching procedural skills and documenting competence.³ The Accreditation Council for Graduate Medical Education (ACGME) and its review committees have set standards for the number of operative cases for surgery residents,⁴ and the amount of time a family medicine resident must spend in the emergency room,⁵ but a standardized method of assessing procedure competence within specialties has not been specified.

Resident technical skills are often learned through traditional apprenticeship by assisting attending physicians and becoming more independent over time.^{6,7} Evaluation of performance commonly relies on subjective faculty

TABLE 1 SCORES BY GROUP AND SPECIALTY

Specialties	n	Incoming Score, %	Outgoing Score, %	P Value
Nonsurgical residents	65	67.2	73.1	<.001
Family medicine	12	73.3	79.6	.05
Internal medicine	17	66.4	68.8	.54
Neurology	5	61.2	63.2	.89
Pediatrics	12	66.8	72.5	.14
Transitional	19	65.7	75.7	.003
Surgical residents	41	78.4	87.7	<.001
Emergency medicine	16	80.2	89.0	.02
General surgery	8	80.8	88.5	.14
Obstetrics	8	72.4	79.7	.33
Orthopedics	5	75.3	92.2	.08
Otolaryngology	4	81.9	93.8	.27
Total	106	71.5	78.7	<.001

assessments, which are often unreliable with little agreement among 2 faculty evaluating the same resident.⁸ The ACGME may set a minimum number of procedures or emergency room hours a resident must complete, but programs must develop reliable and valid methods to assess technical ability and ensure competence. In obstetrics, an objective, structured assessment has been shown to be a valid and reliable method for assessing resident technical skills.^{9–12} The purpose of this study was to assess laceration management performance objectively among surgical and nonsurgical, postgraduate year-1 (PGY-1) residents and to test for interval improvement.

Methods

Participants

This study evaluated 106 PGY-1 residents from 10 medical specialties before and after internship from 2006 to 2008 with a simulated surgical skills station using pigs' feet. Residents were classified as surgical (emergency medicine, general surgery, obstetrics and gynecology, orthopedics, and otolaryngology) and nonsurgical (family medicine, internal medicine, neurology, pediatrics, and transitional year). Emergency medicine residents were included in the surgical cohort because they have a level of exposure to laceration repair training similar to that of surgical residents. To be included in the study, residents had to have completed their entire PGY-1 year of residency training and have completed the simulated surgical skills station at the beginning and end of the PGY-1 year. All residents completed their required

rotations for the first year of their specialty training before retesting.

Surgical Skills Station Scenario

Residents participating in the study were given 11 minutes to manage a 19-year-old man who was stabbed in the right thigh and was in considerable pain with some wound contamination. The residents were tasked with repairing a deep, 3-cm laceration extending through the muscle fascia over the vastus lateralis. Instruments were already laid out on a sterile tray with a pig's foot on a gurney. A sterile drape and cup of iodine swabs were available on the tray. The resident had to select an appropriate suture for deep layer and skin closure; place the suture on the tray in a sterile manner; don sterile gloves; describe the techniques for administering local anesthetic, irrigating, and debriding the wound; and demonstrate closure of the deep and superficial layers. Residents were also asked questions regarding wound irrigation, anesthetic, postprocedural care, and management of postprocedural complications from a standardized script.

The simulated scenario and evaluation form were created by a board-certified general surgeon and reviewed for accuracy and content by 4 additional physician educators. The station was standardized to ensure that all residents received the same information without hints, reminders, or recommendations from the evaluator. The timing of each station was strictly enforced.

Assessment Tool Reliability and Validity

The reliability of the assessment tool was strengthened by providing standardized instructions and education to

TABLE 2 SURGICAL SKILLS STATION CHECKLIST RESULTS BEFORE AND AFTER POSTGRADUATE YEAR-1

Evaluation Form Questions and Answers	All Residents (N = 106)		Surgical Residents (n = 41)		Nonsurgical Residents (n = 65)	
	Before, %	After, %	Before, %	After, %	Before, %	After, %
Select appropriate suture for deep layer and skin closure						
Selects proper suture for deep layer	61.3	94.3	68.3	97.6	56.9	92.3
Selects proper suture for skin closure	65.1	91.5	78	100	56.9	86.2
Patient preparation						
Uses sterile technique to place suture on tray	93.4	97.2	97.6	95.1	90.8	98.5
Uses sterile technique to don gloves	96.2	96.2	95.1	95.1	96.9	96.9
Uses povidone-iodine for sterile prep	100	96.2	100	95.1	100	96.9
Uses all 3 povidone-iodine swabs	92.5	95.3	92.7	95.1	92.3	95.4
Uses swabs in circular motion from wound edges outwards	92.5	93.4	95.1	92.7	90.8	93.8
Applies sterile drape over laceration	90.6	94.3	95.1	92.7	87.7	95.4
Appropriately loads needle on needle driver	84.9	82.1	92.7	90.2	80	76.9
Uses tissue forceps for repair	82.1	83	95.1	92.7	73.8	76.9
What would you use to irrigate the wound?						
Normal saline	97.2	95.3	97.6	100	96.9	92.3
How would you irrigate the wound?						
Normal saline, 1-L bottle with pressure	93.4	77.4	92.7	92.7	93.8	67.7
Would you give anesthetic before or after cleaning the wound? Why?						
Before, because the patient is in considerable pain	65.1	58.5	68.3	70.7	63.1	50.8
Would you inject anesthetic through the wound or through the skin surface? Why?						
Through wound, because this would further decrease pain	55.7	50	58.5	63.4	53.8	41.5
What anesthetic would you use?						
Lidocaine or lidocaine with epinephrine	95.3	98.1	95.1	95.1	95.4	100
When would you NOT use anesthetic with epinephrine?						
Ear	71.7	73.6	78	85.4	67.7	66.2
Nose	81.1	87.7	92.7	97.6	73.8	81.5
Digits	84.9	98.1	95.1	100	78.5	96.9
Penis	71.7	84	82.9	90.2	64.6	80
Why do you NOT use epinephrine in these areas?						
Causes vasoconstriction and possible ischemia	88.7	99.1	95.1	100	84.6	98.5

TABLE 2 CONTINUED

Evaluation Form Questions and Answers	All Residents (N = 106)		Surgical Residents (n = 41)		Nonsurgical Residents (n = 65)	
	Before, %	After, %	Before, %	After, %	Before, %	After, %
What are the advantages of using anesthetic with epinephrine?						
Decreases bleeding due to vasoconstriction	95.3	98.1	95.1	97.6	95.4	98.5
Keeps lidocaine in wound longer	43.4	72.6	61	92.7	32.3	60
Would you place a dressing on the wound after closure or leave open to the air?						
Place dressing	84	81.1	90.2	92.7	80	73.8
When should the patient remove the dressing?						
2 d	30.2	63.2	34.1	70.7	27.7	58.5
When can the patient shower after this laceration repair?						
2 d	41.5	59.4	51.2	65.9	35.4	55.4
In how many days would you remove the skin sutures?						
7–14 d	86.8	86.8	92.7	85.4	83.1	87.7
In how many days would you remove the deep sutures?						
Never, absorbable sutures	79.2	85.8	80.5	85.4	78.5	86.2
What would you do if the patient returns to your clinic in 5 d with evidence of a wound infection and pus coming from the wound?						
Open wound	83	94.3	95.1	97.6	75.4	92.3
Was sterile technique maintained throughout the entire procedure?						
Yes	78.3	92.5	87.8	100	72.3	87.7
Demonstrates proper technique for laceration repair	Average Likert Score		Average Likert Score		Average Likert Score	
Uses proper technique to drive needle	2.43	2.82	2.88	3.95	2.15	2.11
Proper technique for fascial closure with a simple stitch	2.39	2.81	2.83	3.98	2.11	2.08
Proper technique for an instrument tie	2.55	2.91	3.22	4.15	2.12	2.12
Proper technique for simple interrupted stitch	2.45	2.95	2.98	4.20	2.12	2.17
Proper technique for hand tie	2.24	2.58	2.76	3.71	1.91	1.86
Proper technique for interrupted vertical mattress stitch	1.31	2.09	2.07	3.41	0.83	1.26
Proper technique for interrupted horizontal mattress stitch	1.07	1.59	1.63	3.05	0.71	0.68

evaluators on how to use the evaluation forms and by using multiple checklist items, most of which contained objective “yes/no” selections. Subjective checklist items regarding technique relied on the evaluators’ internal criteria for what was considered average for technical skills. The practice effect bias from using the same scenario for testing before and after PGY-1 was minimized by not providing residents feedback after their initial test and by not informing residents that the second test would use the same scenario. This bias was also decreased by the time interval between testing.¹³

The validation of the assessment tool used in this study focused on the 5 sources of evidence used to support construct validity.¹⁴ The content used in this station was researched, and the evaluation form was created by a qualified general surgeon with peer review from other specialty program directors. The response process was improved by using an electronic evaluation form, educating evaluators on the use of the form, and allowing educators to practice using the form and ask questions. For relation to variables, validity was supported by an expected initial score discrepancy between surgical and nonsurgical residents, improved performance over time on subsequent testing, and limited improvement among specialties that do little suturing during the PGY-1 year.

Evaluation and Data Interpretation

One of 2 faculty physicians evaluated the residents before and after internship using a computerized evaluation form with both objective and subjective criteria. The station was evaluated by one general surgeon during the first academic year and by another surgeon the second academic year. The same evaluator was used for the beginning and end of year assessments. An objective checklist, graded in a binomial (yes/no) format was used for suture selection, wound preparation, and questions regarding wound irrigation, anesthetic, postprocedural care, and management of postprocedural complications. Wound closure technique was graded using a 5-point Likert scale (1, needs improvement; 2, below average; 3, average; 4, above average; and 5, excellent). The scores were reported as a percentage of the maximum 180 points available.

Data were collected on standardized InfoPath forms (Microsoft Corporation, Redmond, WA) and consolidated into an Excel spreadsheet (Microsoft). Statistical analysis was done using SPSS version 16 (formerly SPSS Inc, now IBM Corporation, Armonk, NY). Preinternship and postinternship scores were compared by residency department, surgical and nonsurgical groups, and overall, with a paired *t* test and Wilcoxon signed rank test for small sample sizes. A Mann-Whitney *U* test and a χ^2 test were used to compare demographics between the surgical and nonsurgical groups.

This study received institutional review board approval from the Madigan Healthcare System, Fort Lewis, Washington.

Results

A total of 116 PGY residents from 10 residency programs participated in the incoming surgical skills station in June of 2006 and 2007. Ten residents (9%) did not participate in the end-of-year scenarios because of scheduling conflicts. The data analysis for this study was based on the 106 examinees who participated in both the incoming and outgoing surgical skills station.

Participants ranged in age from 25 to 44 years (average, 28.7 years). There were more men (*n* = 70, 66%) than women (*n* = 36, 34%), and more allopathic-MDs (*n* = 75, 71%) than osteopathic-DOs (*n* = 31, 29%) residents, and all participants were graduates of US medical schools. There were 41 surgical residents (mean age, 28.5 years; 29 men [71%]; 36 MDs [88%]) and 65 nonsurgical residents (mean age, 28.8 years; 41 men [63%]; 39 MDs [60%]). The ages (*P* = .23) and sex (*P* = .55) of the residents in the surgical and nonsurgical groups were similar; there were more MDs in the surgical group (*P* = .002). The represented programs included internal medicine (*n* = 17, 16%), transitional year (*n* = 19, 18%), emergency medicine (*n* = 16, 15%), family medicine (*n* = 12, 11%), pediatrics (*n* = 12, 11%), general surgery (*n* = 8, 7.5%), obstetrics and gynecology (*n* = 8, 7.5%), neurology (*n* = 5, 5%), orthopedics (*n* = 5, 5%), and otolaryngology (*n* = 4, 4%).

Surgical residents outscored nonsurgical residents before (78.4% versus 67.2%, *P* < .001) and after (87.7% versus 73.1%, *P* < .001) internship. Scores significantly improved in both groups at the end of the academic year, with improvement seen in all specialties (TABLE 1). Overall, residents showed the most improvement on the topics of suture selection, use of lidocaine with epinephrine, and postprocedural wound care. Subjective scoring of laceration technique, however, improved in the surgical group, but not in the nonsurgical group after the PGY-1 year (TABLE 2). There was no statistically significant difference between allopathic and osteopathic residents on the initial or end of year assessment.

Both the surgical and nonsurgical groups had statistically significant improvement in their scores at the end of the PGY-1 year. Interestingly, the amount of improvement was similar between the 2 groups (surgical, 9.4%; nonsurgical, 5.9%). The evaluation form was designed to test basic laceration skills and knowledge. Many of the residents in the surgical group possessed most of this knowledge after medical school. Surgical scores near 80% on the initial assessment left little room for improvement despite specific training in laceration management. Nonsurgical incoming scores in the mid-60% range left substantial room for improvement.

Group comparisons on actual laceration repair technique did show substantial improvements in Likert scores in the surgical group compared with worsening scores in nonsurgical residents after internship (TABLE 2). This improvement was statistically significant and likely due to additional suture technique training in surgical programs.

This finding is important because it also helps to support the validity of using a surgical skills assessment to show improvement and to determine level of expertise.

Discussion

This study showed interval improvement in both surgical and nonsurgical residents after the PGY-1 year, with all residency programs showing some improvement. Residents in the surgical group had higher scores than the nonsurgical group on both the incoming and outgoing stations. Higher surgical group scores on the incoming assessment were expected given that medical students pursuing surgical training programs tend to spend more time on surgical rotations. The surgical group also outscored the nonsurgical group on the end of year assessment after training in programs that focus on procedural skills.

Limitations of this study include the limitations of the assessment tool, both in the area of internal structure and sources of validity evidence. The internal structure of the assessment tool could have been strengthened by determining reliability coefficients, standard errors of measurements, or κ and weighted- κ testing. The evidence of consequences was not determined in this study, but this source of validity could be strengthened by doing further subgroup analysis to determine performance among those expected to perform similarly. There is potential for interrater bias because 2 different surgeon evaluators were used in the 2 academic years. This bias may exist for Likert scale assessments but is minimal for objective checklist items.

Other potential study limitations include nonblinded evaluators and the simulated nature of the assessment compared with live patients. If this type of station were used for a high-stakes competency assessment, it would also be important to lengthen the station, include resident consent of the patient, and require actual irrigating, anesthetizing, and dressing of the wound.

The validity of using simulation for the assessment and improvement of procedural skills suggests that the time has come to move away from pure traditional apprenticeship as a means of determining procedural competence in residency. Implementation of the ACGME Outcome Project and Milestone Project requires that residents be accurately evaluated and deemed competent in specific skills before advancing to the next level of training. The laceration management assessment used in this study provides a valid framework for the creation of skills qualification checklists for essential procedures within residency programs. High-stakes training checklists need to be validated and to include the full spectrum of procedural care: proper consent, hand-washing, patient preparation, technical skills, and an

understanding of medications, postprocedural care, and wound complications.

Conclusions

Our study demonstrates that measurable differences in laceration management can be demonstrated in the same resident group before and after a year of residency training. It further confirms our hypothesis that those residents with formal training in the procedure should score better than those who do not have such training. To ensure competence in a particular procedure, checklists should include all essential tasks expected of the trainee. The goal in this case is not to detect improvement but to set standard benchmarks that all residents must achieve to be deemed competent. This can be achieved either through specific, Milestone Project working groups or by individual Residency Review Committees.

Future studies should continue to validate qualification checklists to measure competence in basic skills and procedures and to include assessments of interrater reliability by having multiple evaluators assess videotaped sessions.

References

- 1 Singer AJ, Thode HC Jr, Hollander JE. National trends in ED lacerations between 1992 and 2002. *Am J Emerg Med*. 2006;24(2):183–188.
- 2 Croft SJ, Mason S. Are emergency department junior doctors becoming less experienced in performing common practical procedures? *Emerg Med J*. 2007;24(9):657–658.
- 3 Gaies MG, Landrigan CP, Hafner JP, Sandora TJ. Assessing procedural skills training in pediatric residency programs. *Pediatrics*. 2007;120(4):715–722.
- 4 Accreditation Council for Graduate Medical Education. Program requirements for graduate medical education in general surgery. http://www.acgme.org/acWebsite/downloads/RRR_progReq/440_general_surgery_01012008_uo8102008.pdf. Assessed October 15, 2010.
- 5 Accreditation Council for Graduate Medical Education. Program requirements for graduate medical education in family medicine. http://www.acgme.org/acWebsite/downloads/RRR_progReq/120pr07012007.pdf. Assessed October 15, 2010.
- 6 Reichel JL, Peirson RP, Berg D. Teaching and evaluation of surgical skills in dermatology: results of a survey. *Arch Dermatol*. 2004;140(11):1365–1369.
- 7 Mandel LP, Lentz GM, Goff BA. Teaching and evaluating surgical skills. *Obstet Gynecol*. 2000;95(5):783–785.
- 8 Reznick RK. Teaching and testing technical skills. *Am J Surg*. 1993;165(3):358–361.
- 9 Nielsen PE, Foglia LM, Mandel LS, Chow GE. Objective structured assessment of technical skills for episiotomy repair. *Am J Obstet Gynecol*. 2003;189(5):1257–1260.
- 10 Lentz GM, Mandel LS, Goff BA. A six-year study of surgical teaching and skills evaluation for obstetric/gynecologic residents in porcine and inanimate surgical models. *Am J Obstet Gynecol*. 2005;193(6):2056–2061.
- 11 Goff B, Mandel L, Lentz G, et al. Assessment of resident surgical skills: is testing feasible? *Am J Obstet Gynecol*. 2005;192(4):1331–1338.
- 12 Chipman JG, Schmitz CC. Using objective structured assessment of technical skills to evaluate a basic skills simulation curriculum for first-year surgical residents. *J Am Coll Surg*. 2009;209(3):364–370.
- 13 Strauss E, Sherman E, Spreen O. *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. 3rd ed. New York, NY: Oxford University Press; 2006:11.
- 14 Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med*. 2006;119(2):166.e7–e16.