

Reliability of Operative Skill Evaluations: How Late Is Too Late to Give Feedback?

Ricardo J. Bello, MD, MPH; Meredith L. Meyer, MD; Damon S. Cooney, MD, PhD; Gedge D. Rosson, MD; Scott D. Lifchez, MD; Carisa M. Cooney, MPH

INTRODUCTION

Evaluating resident operative skill acquisition is a common challenge for all surgical specialties. The Operative Entrustability Assessment (OEA) is a validated assessment tool that facilitates compliance with the Accreditation Council for Graduate Medical Education's Next Accreditation System and documents resident operative performance at point-of-care.¹ The OEA and other operative rating tools have been implemented in surgical training programs across the United States in an effort to incorporate objective, valid, and reliable operative skills assessments into surgical training.² A recent multi-institutional qualitative study on resident feedback needs raised questions on the reliability of operative skills feedback when this feedback is given more than a week after the case date.³ Using our experience with the OEA, we assessed the reliability of evaluation scores according to the timeliness of feedback completion.

METHODS

We extracted evaluator and self-assessment scores from all logged cases since OEA implementation at our institution. We defined OEA score reliability as the correlation between self-assessment and evaluator scores. This correlation has been shown in previous studies to be positive, moderate to strong, and statistically significant.^{4,5} We used paired *t* test to compare scores and Pearson's correlation coefficient to assess reliability stratifying by time-to-evaluation completion divided into quintiles (Q1: 0, Q2: 1–3, Q3: 4–13, Q4: 14–38, and Q5: > 38 days after surgery). We used likelihood ratio tests on linear regression to assess the interaction between reliability and timeliness to completion.

RESULTS

Between September 2013 and October 2016, 1778 complete OEAs were logged. Mean resident self-assess-

ment (3.41 ± 1.09) was slightly higher than evaluator score (3.37 ± 0.99 ; $P = 0.048$). Overall, self-assessment score was significantly and strongly correlated with evaluator score [Pearson's correlation coefficient ($r = 0.72$; $P < 0.001$)]. Stratified by delay-to-completion, correlation coefficients were roughly similar for evaluations completed within 0 days ($r = 0.77$; $P < 0.001$), 1–3 days ($r = 0.73$; $P < 0.001$), and 4–13 days after surgery ($r = 0.70$; $P < 0.001$). Although still statistically significant, this correlation was only moderately positive for evaluations entered within 14–38 days ($r = 0.60$; $P < 0.001$) or over 38 days ($r = 0.52$; $P < 0.001$) after surgery. We found strong evidence for an interaction between the time to completion of OEA scores and OEA evaluator score reliability ($P < 0.001$).

CONCLUSIONS

Our data support the reliability of OEA evaluator scores completed until 2 weeks from the case, with significantly decreased reliability associated with delayed completion. This represents a useful refinement in the interpretation of evaluation scores that is crucial as we move toward competency-based accreditation in surgical specialties.

Carisa M. Cooney, MPH

Department of Plastic and Reconstructive Surgery
Johns Hopkins University School of Medicine
601 N. Caroline St.
JHOC 8161
Baltimore
MD 21287

E-mail: ccooney3@jhmi.edu

Disclosure: Ms. Carisa Cooney, Dr. Scott Lifchez, and Dr. Damon Cooney are co-inventors of the Operative Entrustability Assessment and co-advisors and equity holders for its parent company, EduMD, LLC, for which Ms. Cooney is the vice president. No other authors have a financial interest in any of the products, devices, or drugs mentioned in this article. The Article Processing Charge for this abstract was paid for by the American Council of Academic Plastic Surgeons.

From the Department of Plastic and Reconstructive Surgery, Johns Hopkins University School of Medicine, Baltimore, Md.

Presented at the American Council of Academic Plastic Surgeons 4th Annual Winter Retreat, December 9–11, 2016, Chicago, Ill.

Copyright © 2017 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of The American Society of Plastic Surgeons. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Plast Reconstr Surg Glob Open 2017;5:e1465; doi:10.1097/GOX.0000000000001465; Published online 25 September 2017.

ACAPS: American Council of Academic Plastic Surgeons (ACAPS) Winter Retreat, in Chicago, Illinois, December 9–11, 2016.

ACKNOWLEDGMENTS

Michael Cohen assisted with data extraction for Operative Entrustability Assessment data. He was not compensated for this contribution. We thank the residents and faculty at the Johns Hopkins School of Medicine Department of Plastic and Reconstructive Surgery in making use of and continuously providing feedback on the Operative Entrustability Assessment.

REFERENCES

1. Cooney CM, Cooney DS, Bello RJ, et al. Comprehensive observations of resident evolution: a novel method for assessing procedure-based residency training. *Plast Reconstr Surg.* 2016;137:673–678.
2. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med.* 2006;355:2664–2669.
3. Bello RJ, Sarmiento S, Rosson GD, et al. Understanding the role for operative performance rating tools in meeting surgical trainee feedback needs: a qualitative study. *Plast Reconstr Surg.* 2016;4:e780.
4. Ward M, MacRae H, Schlachta C, et al. Resident self-assessment of operative performance. *Am J Surg.* 2003;185:521–524.
5. Osborne AJ, Hawkins SC, Pournaras DJ, et al. An evaluation of operative self-assessment by UK postgraduate trainees. *Med Teach.* 2014;36:32–37.