Establishing Validity Evidence for an Operative Performance Rating System for Plastic Surgery Residents

Kim A. Bjorklund, MD, FRCSC, MEd,* Nicole Sommer, MD, FACS,† Michael W. Neumeister, MD, FACS, FRCSC,† and Steven J. Kasten, MD, MHPE‡

*Section of Plastic and Reconstructive Surgery, Department of Surgery, Nationwide Children's Hospital, The Ohio State University College of Medicine, Columbus, Ohio; †Department of Surgery, Southern Illinois University School of Medicine, Springfield, Illinois; and ‡Section of Plastic Surgery, Department of Surgery, University of Michigan, Ann Arbor, Michigan

OBJECTIVE: The aim of this study was to describe an operative performance rating system for plastic surgery residents and provide validity evidence for the instrument.

METHODS: Three plastic surgery residents (PGY levels 1, 5, and 6) from Southern Illinois University School of Medicine (SIUSOM) performed a carpal tunnel release with audio video recording. The 3 videos were reviewed by 8 expert hand surgeons and 3 SIUSOM faculty using the operative performance rating system instrument to assess resident operative performance. Validity evidence including content, internal structure, and relationship to other variables was collected.

RESULTS: Inter-rater reliability was consistently fair to moderate (weighted Cohen's Kappa 0.44-0.84 for experts, 0.24-0.55 for SIUSOM raters), and all assessment items were highly correlated (Cronbach's alpha of 0.9867). Local SIUSOM faculty routinely demonstrated higher overall scores for PGY 1 and PGY 6 residents compared to expert raters.

CONCLUSIONS: Although limited by small numbers, this pilot study suggests that potential bias based upon PGY year, identity, and performance history may exist and independent assessment by unbiased raters or comparison to national operative norms may be valuable. Our study provides baseline validity evidence for a resident operative performance assessment tool that can be integrated into practice in plastic surgery training programs. (J Surg Ed 76:529–539. © 2018 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: resident education, assessment, evaluation of assessment, surgical training

COMPETENCIES: Patient Care, Practice-Based Learning and Improvement

INTRODUCTION

Surgical competence requires a number of skill sets, knowledge, judgment, and technical skill, which can be gained through practice, assessment, and feedback. In order to train competent independent surgeons, residency programs must not only provide training for the acquisition of these skills but also measure trainees' success in their skill acquisition. Resident assessment in plastic surgery has traditionally used global assessments from faculty at designated time periods, such as the end of clinical rotations, followed by annual assessments. These assessments make qualitative judgments across a range of skills and competencies over an extended period of time. Studies have demonstrated that these assessments reflect a subjective, generalized impression of clinical performance and professional behavior.[1-4]

In the operating room, many surgeons rely on opinion and experience for resident teaching and assessment, with few having formal training in education.[5] Unfortunately, this lack of standardized assessment of resident performance makes it difficult to objectively assess resident competence in specific procedures and provide timely formative feedback. Although resident case logs provide an estimate of operative experience, they do not directly assess resident operative performance. In an

Correspondence: Inquiries to Kim A. Bjorklund, MD, FRCSC, MEd, Department of Surgery, Section of Plastic and Reconstructive Surgery, Nationwide Children's Hospital, The Ohio State University College of Medicine, 700 Children's Drive, Columbus, OH 43205; e-mail: kimabjorklund@gmail.com

  1931-7204/$30.00 https://doi.org/10.1016/j.jsurg.2018.08.016

age of increasing financial and medicolegal demands with decreased resident work hours,[5] demonstrating the ability to competently perform surgical procedures is an essential component of the patient care competency. The ACGME mandates that plastic surgery training programs provide objective assessments of trainee competence in patient care and procedural skills in order to complete the program.[6]

The operative performance rating system (OPRS), developed by the Department of Surgery at Southern Illinois University School of Medicine (SIUSOM), consists of procedure-specific assessments of resident operative performance. The OPRS consists of a skill-specific checklist, general skill assessment, and degree of prompting by the attending surgeon.[4] The instrument is widely used in general surgery at SIUSOM and has been developed and validated for the subspecialties of urology and laparoscopic surgery.[7,8] At present, no procedure-specific intraoperative resident assessment for which there is good validity evidence is routinely used in practice across plastic surgery training programs in the United States. Van Heest et al. developed and provided validity evidence for orthopedic residents performing carpal tunnel releases (CTR) on cadavers in the skills lab using the Objective Structured Assessment of Technical Skills.[9,10] CTR is also a common procedure performed in plastic surgery residency. It is a foundational skill noted in the ACGME milestones and one of the most frequently performed procedures nationally, according to case log data.[11,12]

The purpose of this study is to develop a CTR OPRS assessment for plastic surgery residents and provide validity evidence consistent with the *Standards for Educational and Psychological Testing,* developed jointly by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education.[13] The *Standards* provide collaborative benchmarks for testing and assessment practices.[13] Building upon the OPRS framework and procedure-specific CTR performance assessments conducted in the cadaver lab,[14,15] we aim to provide validity evidence for this tool upon implementation in the operating room with plastic surgery residents. The intended use of the assessment tool is to direct the observations of surgeons when judging resident operative performance and to use those judgments to provide an indication of a resident's ability to perform the procedure independently.

METHODS

Institutional Review Board exemption was obtained from the Springfield Committee Involving Human Subjects in Research. Three separate audio video recordings of integrated plastic surgery residents in years PGY 1, 5, and 6 performing a CTR procedure were completed. The 3 videos were subsequently reviewed by a group of expert hand surgeons and by SIUSOM faculty using the OPRS instrument to assess resident operative performance 3 to 4 months after the procedures were performed and recorded.

Audio and/or Video Records of Carpal Tunnel Release Procedures

Three residents (PGY1, 5, and 6) were each evaluated performing a separate CTR procedure in the operating room using a head-mounted GoPro 3+Silver camera worn by the resident. In each of the 3 videos, the resident performed the procedure with an attending hand surgeon present. Residents were selected based upon PGY year, with the goal of having a junior resident, chief resident, and mid-level or senior resident. Based upon the timing of resident rotations during the study, a sample of convenience with a PGY 5 at the beginning of the fifth year was employed as the senior mid-level resident.

Videos were edited to focus on key steps of the procedure and remove patient identifying information. The videos ranged from 6 to 8 minutes in length after editing from the initial recording time of 10 to 15 minutes per video. The editing process took approximately 1 to 2 hours per video. Audio was included to help in the assessment of resident performance and identify the amount of guidance required by the attending surgeon. The camera view focused on the operative field throughout the procedure and included the hands of the resident, supervising attending surgeon, and surgical tech.

Participants

Eight expert surgeons were selected to review the same 3 audio video records of resident CTR procedures using the CTR OPRS instrument. The number of surgeons was selected based on standard setting research indicating that samples of 7 experts provide a good estimate of the consensus of the population of expert judges.[16,17] Selected expert surgeons were plastic or orthopedic surgeons who routinely performed CTR in practice, supervised residents performing CTR, and did not know the residents in the study. Participating surgeons were from 5 different residency training programs across the United States. Hand surgery fellowships had been completed by 7 out of 8 experts, and all experts had extensive experience performing CTR in practice.

Three faculty hand surgeons from SIUSOM plastic surgery program were also asked to review the 3 audio video recordings. These 3 surgeons routinely performed CTR in practice, educated, and supervised the SIUSOM

plastic surgery residents performing CTR and knew all residents in the study. Two of the 3 surgeons were also the supervising surgeons in the video recordings.

Operative Performance Rating Instrument

The OPRS for CTR was developed using the format of SIUSOM general surgery OPRS checklists. CTR was chosen as a sentinel procedure based upon review of the ACGME Milestones and resident logs of the most commonly performed plastic surgery procedures, both at SIUSOM and nationally.

The assessment tool (Fig. 1) included 5 procedure-specific items, 4 general items, and an overall performance item. Each procedure-specific item was rated for resident performance and degree of attending surgeon guidance. Additional feedback was solicited at the end of the assessment with comment boxes regarding the resident's strengths and weaknesses. General items were derived from items developed by the University of Toronto and include instrument handling, respect for tissue, time and motion, and operation flow.[18] Items for procedure-specific steps were developed by reviewing the literature for the key steps that affect procedural outcomes in CTR. This included technique-specific articles[14,15] as well as previous literature highlighting the key steps of CTR for skill assessment in trainees.[9,10] These items were subsequently reviewed by experts in the field for refinement. A 5-point Likert scale with descriptive anchors was used for each item (poor, fair, good, very good, and excellent) as well as a "not applicable" option if the resident did not perform the step. The degree of guidance required from the attending surgeon was noted for each step on a 1 (substantial prompting required) to 5 (minimal prompting required) scale. Finally, the overall difficulty of the particular case was assessed using a 5-point Likert scale ranging from 1 (straightforward) to 5 (very difficult). Case difficulty may be increased by factors such as abnormal anatomy, previous surgery in the area, or patient characteristics such as morbid obesity.

Rating Process

All participating surgeon raters independently reviewed the 3 audio video CTR recordings and rated each resident performance using the CTR OPRS instrument in paper form. Instructions indicated that the OPRS assessment for each resident should be completed immediately after observing their video-recorded operative performance. No restrictions or limitations were given regarding the number of times a performance could be observed or if raters wished to pause and replay specific aspects of a performance.

Raters were provided with an overview of the study and instructions to guide the use of the OPRS instrument. All surgeons were assured that their assessments were confidential, for research purposes only, and would not be provided to the residents or have any impact on resident progress. The expert rater surgeons were blinded to the identity, PGY level, and performance history of all the residents.

DATA ANALYSIS

Analysis was performed using SAS 9.3 (SAS Institute, Cary, NC). Absolute (simple kappa) and relative (weighted kappa) agreement was used to assess inter-rater reliability. Wilcoxon rank-sum test, the nonparametric method, was used to test the difference between expert group and SIU surgeon group. Cronbach's alpha was used to estimate the interitem correlation between general and specific OPRS items and the internal consistency. Mean OPRS scores and standard deviations for the residents were computed.

RESULTS

Table 1 provides a summary of overall performance ratings for each resident, including mean expert ratings and standard deviations overall and for each item within the OPRS assessment. All expert raters deemed the cases to be low in difficulty.

Comparison of expert scoring among 3 residents for each item indicated the PGY 5 resident with 54 CTR procedures recorded in the case log obtained the highest scores from experts in each item, while the PGY 1 resident with only 2 CTR procedures had the lowest scoring. The PGY 6 resident with 84 CTR procedures logged was scored in the middle. Expert raters noted that significant overall direction (1.13) was required by the PGY 1 resident, while some direction (3.25) was required by the PGY 6 resident and minimal direction (4.63) by the PGY 5 resident.

Internal Consistency Reliabilities

Internal consistency reliability for general and specific items was high with a Cronbach's alpha of 0.9867, indicating that items are reflective of the same underlying construct.

Inter-rater Reliability

Inter-rater reliability was calculated using simple and weighted Cohen's kappa for both the SIUSOM surgeon raters and expert raters. All 3 video assessment scores were pooled together to calculate the Cohen's kappa.

Open Carpal Tunnel Release - Rating Form

Evaluator:	**Resident:**
Resident Level:	**Program:**

Date of Procedure:	**Time Procedure Was Completed:**
Date Assessment Was Completed:	**Time Assessment Was Initiated:**

Please rate this resident's performance during this operative procedure. The caption above each item provides descriptive anchors for 3 of the 5 points on the rating scale. "N/A" (Not applicable) should only be selected when the resident did not perform that part of the procedure.

OPRS (Operative Performance Rating) Case Difficulty

Indicate the difficulty of the case:

1	2	3	4	5
Straightforward anatomy, no related prior surgeries or treatment		Intermediate difficulty		Abnormal anatomy, prior CTR or surgery in area, morbid obesity
0	0	0	0	0

Degree of Prompting or Direction

Substantial direction. 1	**2**	**Some direction. 3**	**4**	**Minimal direction. 5**
Unable to direct team, use / choose instruments, anticipate next steps as surgeon or as first assistant, without constant attending prompting.		Actively assists & anticipates own & attending needs, performs basic steps with occasional attending direction to resident and /or to surgical team. Somewhat hesitant &		Performs all steps & directs team with minimum direction from attending to resident or to team i.e. anticipates needs, sets up exposure for self & assistant, transitions fluently between
		slow to anticipate or recognize aberrant anatomy, unexpected findings & / or "slowing down moments."		steps, gives clear direction to first assistant, maintains situation awareness, calmly recovers from error & recognizes when to seek help/ advice
0	0	0	0	0

Procedure-Specific Criteria

Please indicate the degree of prompting for **each item.** The assessment score for each item **MAY** differ from the prompting score for that item.

Positioning, Prepping, Draping

1 Poor	2 Fair	3 Good	4 Very Good	5 Excellent	NA
Unfamiliar with appropriate set up such as tourniquet, hand table, appropriate side for surgeon		Adequate but omits or incorrectly performs some essential steps		Safe, efficient and optimal positioning of patient and surgeon for procedure. Directs OR team in prepping and draping	
0	0	0	0	0	0

Substantial Direction 1	**2**	**Some direction. 3**	**4**	**Minimal direction. 5**	**N/A**
0	0	0	0	0	0

Incision marking

FIGURE 1. OPRS CTR assessment instrument.

1 Poor	2 Fair	3 Good	4 Very Good	5 Excellent	NA
Unaware of or unable to correctly identify landmarks (midline or radial border ring finger (3-5mm from thenar crease), approx 1.5cm incision from within 0.5cm distal palmar crease proximally, need for ulnar jog if carrying incision proximally)		Hesitant but adequate placement of incision		Correctly places incision using appropriate landmarks, aware of pitfalls of incorrect placement (palmar cutaneous branch injury, palmar arch, etc)	
0	0	0	0	0	0

Substantial Direction 1	**2**	**Some direction. 3**	**4**	**Minimal direction. 5**	**N/A**
0	0	0	0	0	0

Surgical incision & dissection to TCL

1 Poor	2 Fair	3 Good	4 Very Good	5 Excellent	NA
Unaware of anatomical layers, does not use retractors for visualization of underlying structures		Hesitant but adequate dissection of layers (fat, palmar fascia) & hesitant placement of retractors to facilitate exposure		Expedient dissection down to TCL. Uses appropriate retractors (Senn, Ragnell, small self-retaining) for optimal exposure	

				of TCL	
0	0	0	0	0	0

Substantial Direction 1	**2**	**Some direction. 3**	**4**	**Minimal direction. 5**	**N/A**
0	0	0	0	0	0

TCL release

1 Poor	2 Fair	3 Good	4 Very Good	5 Excellent	NA
Unable to identify transverse carpal ligament or to enter carpal tunnel, lacks knowledge of distal TCL landmarks/location palmar arch, Guyon's canal, recurrent motor branch median nerve		Hesitant release of TCL, enters carpal tunnel but tentative in complete release		Expedient release of TCL on its ulnar aspect, recognizes distal fat pad as endpoint for distal release, takes care to avoid injury to median nerve throughout release	
0	0	0	0	0	0

Substantial Direction 1	**2**	**Some direction. 3**	**4**	**Minimal direction. 5**	**N/A**
0	0	0	0	0	0

Release proximal antebrachial fascia

1 Poor	2 Fair	3 Good	4 Very Good	5 Excellent	NA

FIGURE 1. Continued.

Does not orient self or assistant at end of table in appropriate position for release. Unfamiliar with need for release with scissors turned in ulnar direction and for need to cut fascia under direct visualization		Hesitant in positioning of self/assistant for release. Uses correct technique but release not complete		Efficient release with correct technique from end of table or directs assistant in complete release using retractor in apex of proximal incision to allow direct visualization of fascia	
0	0	0	0	0	0

Substantial Direction 1	**2**	**Some direction. 3**	**4**	**Minimal direction. 5**	**N/A**
0	0	0	0	0	0

General Criteria

Instrument Handling

1 Poor	2 Fair	3 Good	4 Very Good	5 Excellent
Tentative or awkward movements or inappropriate instrument use.		Competent use of instruments occasionally appeared awkward		Fluid movements with instruments, no stiffness or awkwardness
0	0	0	0	0

Respect for Tissue

1 Poor	2 Fair	3 Good	4 Very Good	5 Excellent
Frequent unnecessary tissue force or damage by inappropriate instrument use.		Careful tissue handling, *occasional* inadvertent damage		*Consistently* handled tissue carefully (appropriately), minimal tissue damage
0	0	0	0	0

Time and Motion

1 Poor	2 Fair	3 Good	4 Very Good	5 Excellent
Many unnecessary moves		Efficient time & motion, some unnecessary moves		Clear economy of motion, and maximum efficiency
0	0	0	0	0

Operation Flow

1 Poor	2 Fair	3 Good	4 Very Good	5 Excellent
Frequent lack of forward progression; frequently stopped operating and		Some forward planning, reasonable procedure progression		Obviously planned course of operation and anticipation of next steps.

FIGURE 1. Continued.

seemed unsure of next move				
0	0	0	0	0

Overall Performance

(NOT included in calculation of mean score) Rating of very good or higher indicates technically proficient performance (i.e., resident is ready to perform operation independently, assuming resident consistently performs at this level)

Poor	Fair	Good	Very good	Excellent
0	0	0	0	0

Please indicate the weaknesses in this resident's performance:

Please indicate the strengths in this resident's performance:

FIGURE 1. Continued.

TABLE 1. OPRS Score Mean and Standard Deviation for Each Resident

	Expert Scoring					
Item	**(PGY6)**	**Standard Deviation**	**(PGY5)**	**Standard Deviation**	**(PGY1)**	**Standard Deviation**
Difficulty	1.00	0.00	1.00	0.00	1.00	0.00
Direction	3.25	0.46	4.63	0.48	1.13	0.33
Positioning	4.00	1.15	5.00	0.00	3.00	1.79
Direction	4.00	1.15	5.00	0.00	3.00	1.79
Marking	3.38	1.19	4.88	0.33	1.13	0.33
Direction	3.13	1.13	5.00	0.00	1.00	0.00
Incision and/or dissection	3.75	0.71	4.88	0.33	2.13	0.78
Direction	3.25	0.46	4.88	0.33	1.63	0.70
TCL release	3.50	0.93	4.50	0.50	2.25	0.66
Direction	3.38	0.92	4.63	0.48	1.63	0.48
Proximal release	3.25	1.04	4.88	0.33	2.13	0.78
Direction	3.38	1.06	4.75	0.43	1.63	0.70
Instrument handling	3.25	0.89	4.75	0.43	1.75	0.66
Tissue	3.50	0.76	4.75	0.43	2.38	1.11
Time and/or motion	2.88	0.99	4.25	0.43	1.75	0.66
Flow	3.25	0.71	4.63	0.48	1.75	0.66
Overall performance	**3.13**	0.83	**4.75**	0.43	**1.50**	0.50

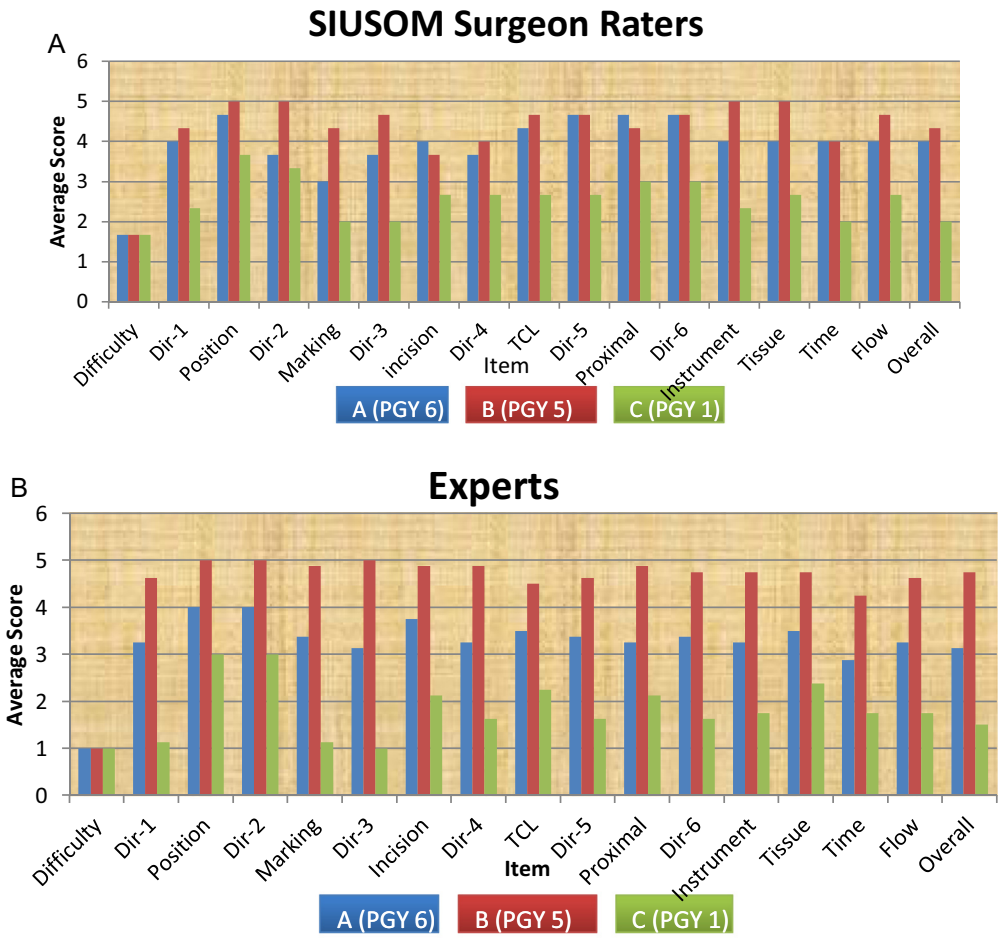

FIGURE 2. A: Comparison of SIUSOM rater scoring across different items. B: Comparison of expert rater scoring across different items.

TheagreementamongtheSIUSOMsurgeonswasnotedtobe fair to moderate (0.24-0.55), while among expert raters noted to be moderate to substantial (range weighted Cohen's Kappa 0.44-0.84).

In examination of overall performance scores, SIUSOM surgeon raters gave scores of 2 (PGY 1), 4 (PGY 5), and 4 (PGY 6), while expert raters gave scores of 1.5 (PGY 1), 5 (PGY 5), and 3 (PGY 6) (Fig. 2 A,B). An overall ranking of good (3) implies a weak performance, while very good (4) is thought to reflect a very satisfactory performance.[18]

DISCUSSION

To maximize patient safety and ensure residents are prepared for independent practice, it is vital to assess residents' ability to perform specific procedures in the operating room. Based upon successful outcomes of OPRS trainee assessments in general surgery,[4,18] laparoscopic surgery,[8] and urology,[7] we applied the OPRS format to the intraoperative assessment of plastic surgery residents. Noland's study of essential hand procedures for mastery by graduating plastic surgery residents indicates that future steps must focus on incorporating structured assessment of these designated procedures into plastic surgery training programs.[19] As open CTR is a critical hand surgery procedure for mastery, we focused on designing and gathering validity evidence for an OPRS resident assessment of this procedure.

Validity evidence to support the assessment was provided within the categories of content, internal structure, and relationship to other variables.[13,20] The content of the assessment was derived from reviewing the literature for previous CTR assessment tools, key elements critical to the successful outcome of the CTR procedure, and refined based upon input from expert surgeons.[9,10,14,15] Both general and specific items were highly correlated and all items had excellent internal consistency. Williams et al. have suggested that using 5 to 7 different expert judges can

reasonably account for individual judge idiosyncrasies, which is supported by the findings in our study.[18]

The inter-rater reliability among expert judges was consistently moderate to substantial, suggesting the overall level of expert rater agreement on resident performance was high. For the local faculty, 2 out of 3 pairings demonstrated moderate agreement and 1 pair demonstrated fair agreement. This may be attributed to a halo and horns effect or recency effect where preconceived opinions or recent trainee performance may have influenced ratings by local faculty.[21] Stringency-leniency error and range restriction may also have affected scoring. Strategies to reduce bias include rater error training to improve awareness of types of rater error, frame of reference training to help calibrate raters, and behavioral observation training to help raters identify critical moments throughout the procedure for assessing the relevant skills.[21] Ongoing feedback and education should be provided to raters over time to identify and correct biases and improve assessment practices. Although limited by the small number of resident performances, our results indicate that SIUSOM faculty routinely scored the PGY 1 and PGY 6 residents higher than expert raters in nearly all assessment items. While the PGY 5 resident clearly scored highest overall in all items by the expert raters, this was not the case with the SIUSOM faculty raters. One explanation for the scoring differences is the inherent bias from faculty raters based upon the residents' PGY year and performance history. Furthermore, the SIUSOM faculty raters maintain an ongoing relationship with the residents and may have felt uncomfortable scoring a PGY 5 resident higher than the PGY 6. These findings are consistent with those of Williams et al. who noted that OPRS assessments of video-recorded general surgery resident procedures yielded supervising surgeon ratings which were higher than expert judge ratings.[18] Other explanations include that the instrument may be more suited to identifying broad extremes of resident performance and yield similar scores once a resident is performing a skill at or beyond a minimum level, which would have been expected for both the PGY 5 and 6 residents. As Bello et al. noted, operative entrustability may be associated with factors unrelated to PGY year.[23] Finally, variability in attending surgeon guidance may also have influenced resident OPRS scores. In this study, the PGY 6 and PGY 1 residents worked with the same attending surgeon during the taped procedure, while the PGY 5 resident worked with a different attending surgeon. Differences in amount and nature of trainee guidance during the procedures may have influenced how expert raters perceived trainee operative performance. This weakness could be addressed by repeating the study using the same attending surgeon for all trainee videos and training faculty to standardize the amount of guidance while maintaining patient safety. Using the instrument in practice to repeatedly assess multiple residents at different PGY levels over time will provide additional information regarding the relationship of resident performance to PGY year and further validate the instrument.

In an era of significant demands on both faculty and trainees' time, it is important to address the feasibility of implementing this procedure-specific assessment tool in practice. According to Williams et al., 20 observations of resident operative performance per year are sufficient as a stable estimate for annual progress decisions.[22] Less than 2 observations per month using the OPRS should be feasible for faculty to build into a rotation. Our experience using this OPRS suggests that it typically takes less than 5 minutes to complete. This is supported by the study of OPRS in general surgery where faculty took an average of 134 seconds to enter 1 trainee assessment using a mobile Web platform.[23] To minimize challenges associated with timing conflicts, OPRS completion may be incorporated into the attending surgeon's postoperative routine (such as performing at the time of operative dictation). Utilizing a mobile app for OPRS could provide reminders and alerts to further encourage timely assessment completion within 72 hours following the procedure. Furthermore, combining observations across a variety of operative procedures using a variety of raters is important in making annual progress decisions with OPRS.[22] Employing these guidelines should help reduce sampling error. To reduce the risk of bias from the Hawthorne effect, OPRS assessments could be performed randomly each month to avoid the effect of observation on the residents' performance. Completion of the assessment after the procedure and without video recording would minimize the Hawthorne effect and should be correlated with additional measures of resident performance when used to determine procedural competence and trainee advancement.

Ultimately, comprehensive resident evaluation benefits from a combination of validated assessment tools. For brief global assessments on a daily basis, generalized instruments such as the system for improving and measuring procedural learning app, comprehensive observations of resident evolution, Ottawa surgical competence operating room evaluation, or operative entrustability assessment may provide rapid generalized feedback. These tools provide succinct surgical assessment that can increase resident and faculty participation and may be used routinely across all procedures.[24-27] However, OPRS can be used to supplement these general tools by providing a deep dive into resident performance for specific procedures. A score above or below a designated threshold on the generalized assessment tools could trigger an OPRS for an in-depth

examination of operative performance and provide specific feedback tailored to a resident's needs. Trialing OPRS alongside existing intraoperative assessment tools in practice could help determine the most effective use of such tools, and whether best applied concurrently or in sequence. While further study is needed to evaluate the utility of an OPRS for plastic surgery, advantages include well-documented reliability, validity, and feasibility in other surgical specialties,[4,7,18] as well as detailed procedure-specific items that may help identify specific performance issues and guide remediation.[28] OPRS may be used in its current form to help plastic surgery programs fulfill their ACGME requirement of objective assessment of trainee procedural skills.[6]

Future directions include implementing the OPRS assessment tool into resident training through an internet-based platform or smart phone app, as well as developing additional OPRS assessments for other sentinel plastic surgery procedures. Finally, directing focus toward an overall level of autonomy that may be tied to the milestones may help make the OPRS assessment more meaningful.[24,29] While not practical to routinely video residents and have them assessed by blinded experts, video recording may be a useful adjunct in the assessment of surgical competence and as a surgical coaching tool to augment resident education.[18,30]

CONCLUSION

In conclusion, the literature reveals a paucity of plastic surgery-based attention to intraoperative resident skill assessment. Over the past 10 years, other surgical specialties have provided significant insight into well-developed intraoperative resident assessments. We have designed and provided validity evidence in the form of content, relationship to other variables, response process, and internal structure for a CTR OPRS for plastic surgery residents as a starting point for development of an intraoperative assessment system in our specialty. Our study provides baseline evidence to begin using this instrument in practice, where ongoing validity evidence should be gathered and used to modify the instrument over time.

REFERENCES

1. Verhulst SJ, Colliver JA, Paiva RE, et al. A factor analysis study of performance of first-year residents. *J Med Educ*. 1986;61:132–134.
2. Ramsay PG, Wenrich MD, Carline JD, et al. Use of peer ratings to evaluate physician performance. *JAMA*. 1993;269:1655–1660.
3. Williams RG, Klamen DA, McGaghie WC. Cognitive, social and environmental sources of bias in clinical competence ratings. *Teach Learn Med*. 2003;15:270–292.
4. Larson JL, Williams RG, Ketchum J, Boehler ML, Dunnington GL. Feasibility, reliability and validity of an operative performance rating system for evaluating surgery residents. *Surgery*. 2005;138:640–649.
5. Kenton K. How to teach and evaluate learners in the operating room. *Obstet Gynecol Clin North Am*. 2006;33:325–332.
6. Accreditation Council for Graduate Medical Education; American Board of Plastic Surgery. Program Requirements for Graduate Medical Education in Plastic Surgery. https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/360_plastic_surgery_2017-07-01.pdf. Accessed June 20, 2018.
7. Benson A, Markwell S, Kohler TS, Tarter TH. An operative performance rating system for urology residents. *J Urol*. 2012;188:1877–1882.
8. Gumbs AA, Hogle NJ, Fowler DL. Evaluation of resident laparoscopic performance using global operative assessment of laparoscopic skills. *J Am Coll Surg*. 2007;204:308–313.
9. Van Heest A, Putnam M, Agel A, Shanedling. J, McPherson S, Schmitz C, et al. Assessment of technical skills of orthopaedic surgery residents performing open carpal tunnel release surgery. *JBJS*. 2009;91:2811–2817.
10. Van Heest A, Kuzel B, Agel J, Putnam M, Kalliainen L, Fletcher J. Objective structured assessment of technical skill in upper extremity surgery. *J Hand Surg Am*. 2012;37:332–337. 337.e1-4.
11. Accreditation Council for Graduate Medical Education; American Board of Plastic Surgery. The Plastic Surgery Milestone Project. https://www.acgme.org/Portals/0/PDFs/Milestones/PlasticSurgeryMilestones.pdf?ver=2016-05-31-113719-180. Accessed August 13, 2018.
12. Accreditation Council for Graduate Medical Education; American Board of Plastic Surgery. ADS/Case Logs Accessed August 13 http://www.acgme.org/Specialties/Case-Log-Information/pfcatid/24/Surgery. Accessed August 13.
13. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. American Educational Research Association, 1999.

14. Nigst H. The carpal tunnel syndrome. operative technique for surgical decompression. *Orthop Traumatol*. 1992;1:122–129.

15. Taleisnik J. The palmar cutaneous branch of the median nerve and the approach to the carpal tunnel: an anatomical study. *JBJS*. 1973;55:1212–1217.

16. Williams RG, Dunnington GL, Klamen DL. Forecasting residents' performance – partly cloudy. *Acad Med*. 2005;80:415–422.

17. Norcini JJ. Setting standards on educational tests. *Med Educ*. 2003;37:464–469.

18. Williams RG, Chen X, Sanfey H, Dunnington G. A controlled study to determine measurement conditions necessary for a reliable and valid operative performance assessment: a controlled prospective observational study. *Ann Surg*. 2012;256:177–187.

19. Noland SS, Fischer LH, Lee GK, Friedrich JB, Hentz VR, et al. Essential hand surgery procedures for mastery by graduating plastic surgery residents: a survey of program directors. *Plast Reconstr Surg*. 2013;132:977e–984e.

20. Beckman TJ, Cook DA, Mandrekar JN. What is the validity evidence for assessments of clinical teaching. *J Gen Intern Med*. 2005;20:1159–1164.

21. Moshe Feldman, et al. Rater training to support high–stakes simulation–based assessments. *J Contin Educ Health Prof*. 2012;32:279–286.

22. Williams RG, Verhulst SJ, Colliver JA, Chen X, Sanfey H, Dunnington G. A template for reliable assessment of resident operative performance: assessment intervals, number of cases and raters. *Surgery*. 2012;152:517–527.

23. Wagner JP, Chen DC, Donahue TR, Quach C, Hines OJ, Hiatt JR, Tillou A, et al. Assessment of resident operative performance using a real-time mobile web system: preparing for the milestone age. *J Surg Educ*. 2014;71:e41–e46.

24. Bello RJ, Major MR, Cooney DS, Rosson GD, Lifchez SD, Cooney CM, et al. Empirical validation of the operative entrustability assessment using resident performance in autologous breast reconstruction and hand surgery. *Am J Surg*. 2017;213:227–232.

25. Cooney CM, Cooney DS, Bello RJ, Bojovic B, Redett RJ, Lifchez SD, et al. Comprehensive observations of resident evolution: a novel method for assessing procedure-based residency training. *Plast Reconstr Surg*. 2016;137:673–678.

26. Gofton WT, Dudek NL, Wood TJ, Balaa F, Hamstra SJ. The Ottawa surgical competency operating room evaluation (O-SCORE): a tool to assess surgical competence. *Acad Med*. 2012;87:1401–1407.

27. Bohnen JD, George BC, Williams RG, et al. The feasibility of real-time intraoperative performance assessment with SIMPL (system for improving and measuring procedural learning): early experience from a multi-institutional trial. *J Surg Educ*. 2016;73:e118–e130.

28. Sanfey H, Williams R, Dunnington G. Recognizing residents with a deficiency in operative performance as a step closer to effective remediation. *J Am Coll Surg*. 2013;216:114–122.

29. Cooney CM, Redett III RJ, Dorafshar AH, Zarrabi B, Lifchez SD, et al. Integrating the NAS milestones and handheld technology to improve residency training and assessment. *J Surg Educ*. 2014;71:39–42.

30. Moore MD, Abelson JS, O'Mahoney P, Bagautdinov I, Yeo H, Watkins AC. Using GoPro to give video-assisted operative feedback for surgery residents: a feasibility and utility assessment. *J Surg Educ*. 2018;75:497–502.