

ORIGINAL RESEARCH

Tracking operative autonomy and performance in otolaryngology training using smartphone technology: A single institution pilot study

Jenny X. Chen MD¹  | Elliott Kozin MD¹  | Jordan Bohnen MD² |
Brian George MD, MSC³ | Daniel Deschler MD¹ | Kevin Emerick MD¹ |
Stacey T. Gray MD¹

¹Department of Otolaryngology–Head and Neck Surgery, Massachusetts Eye and Ear/Harvard Medical School, Boston, Massachusetts

²Department of General Surgery, Massachusetts General Hospital, Boston, Massachusetts

³Department of General Surgery, University of Michigan, Ann Arbor, Michigan

Correspondence

Jenny X. Chen, Massachusetts Eye and Ear, 243 Charles Street, Boston, MA 02114.
Email: jenny_chen@meei.harvard.edu

Funding information

2017 AAO-HNSF CORE Resident Research Award

Abstract

Background: In the era of duty hour restrictions, otolaryngology residents may not gain the operative experience necessary to function autonomously by the end of training. This study quantifies residents' autonomy during key indicator cases, defined by the Accreditation Council for Graduate Medical Education.

Study Design: Prospective cohort study.

Methods: Faculty and residents at a large academic institution were surveyed on the surgical autonomy trainees should achieve for otolaryngology key indicator surgeries at each training level. Residents and faculty used the mobile application “System for Improving and Measuring Procedural Learning” (SIMPL) between December 2017 and July 2018 to log trainees' operative autonomy during cases on a validated four-level Zwisch scale, from “show and tell” to “supervision only.”

Results: The study included 40 participants (23 residents and 17 attendings). The survey response rate was 83%. In surveys, residents overestimated the autonomy PGY5 residents should achieve for parotidectomy, rhinoplasty, thyroid/parathyroidectomy, and airway procedures compared with faculty ($P < .05$). Using SIMPL, 833 evaluations were logged of which 253 were paired evaluations for key indicator cases. Comparing survey predictions with actual cases logged in SIMPL, residents and faculty overestimated the autonomy achieved by senior trainees performing mastoidectomy (PGY5, $P < .05$) and ethmoidectomy (PGY4/5, $P < .05$); both felt that senior residents should operate with between “passive help” and “supervision only” whereas residents actually had “passive help.” Residents overestimated their autonomy during rhinoplasty (PGY5, $P = .017$) and parotidectomy (PGY5, $P = .007$) while attendings accurately expected chief residents to have “passive help.”

Conclusions: Resident surgical autonomy varies across otolaryngology procedures. Multicenter studies are needed to elucidate national trends.

This article was presented at the 2019 Triological Society Combined Sections Meeting.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2019 The Authors. *Laryngoscope Investigative Otolaryngology* published by Wiley Periodicals, Inc. on behalf of The Triological Society.

Level of Evidence: 2

KEYWORDS

operative independence, residency, surgical education

1 | INTRODUCTION

The balance between encouraging the development of resident surgical autonomy and providing appropriate supervision remains a key point of discussion for surgical educators and regulating bodies that provide oversight to resident training. Rising concern for patient safety, duty hour restrictions, and increased expectations for faculty supervision, have created a shift in surgical education from the Halstedian model of “see one, do one, teach one” to a more regulated system of supervised training and evaluation.¹ With the time constraints of duty hour restrictions, the educational emphasis has shifted away from the amount of time spent in the hospital to the quality of the training experiences.² Concurrently, rising financial pressures on hospitals may limit trainee participation as previous investigations of operative cases involving trainees in otolaryngology and other specialties suggest that resident teaching is correlated with longer operative times.³⁻⁵ Increasing operative times may also have ramifications for patient safety, particularly in the pediatric population or for patients with medical comorbidities.⁶

Faculty must balance the desire to optimize the care of their current patients with the need to ensure the safety of trainees' future patients by simulating autonomous practice and encouraging the acquisition of surgical competence.⁷ Limited data exist in the otolaryngology literature about the impact of these challenges on the acquisition of resident operative autonomy and the surgical confidence of graduating residents. There is increasing concern, however, that surgical trainees across subspecialties may not receive sufficient training to become confident and competent surgeons by graduation.⁸ In this study, we sought to quantitatively measure otolaryngology resident autonomy and compare this with the expectations of residents and faculty.

2 | MATERIALS AND METHODS

This study was approved by the Massachusetts Eye and Ear Human Studies Committee. Data were collected from a single large academic otolaryngology residency program. One month prior to the initiation of the study, participants were oriented to the smartphone phone application “System for Improving and Measuring Procedural Learning” (SIMPL) at a standardized 1-hour presentation, during which they were familiarized with the interface, the structured questions, and the use of the Zwisch scale for rating operative autonomy. The training session includes video demonstrations of different levels of autonomy in the operating room. In the field of General Surgery, research has shown that a 1 hour session is sufficient for users to reliably learn how to use the Zwisch scale, which has been highly correlated with

other instruments including the Operative Performance Rating System and the Ottawa Surgical Competency Operating Room Evaluation.⁹

2.1 | Survey administration

In the first phase of the study, participants were sent a Qualtrics survey (Provo, Utah) asking them to predict how much autonomy residents at different level of training might expect while performing 14 otolaryngology key indicator cases as determined by the Accreditation Council for Graduate Medical Education (ACGME) on the Zwisch scale (Table 1). Minimum procedure numbers for key indicator cases have been determined by the ACGME and are tracked by all residency programs to ensure adequate operative experience offered over the course of residency training.¹⁰ The full survey is in Supporting Information Appendix A.

2.2 | SIMPL

In the second phase of the study, faculty and residents used SIMPL to assess resident performance in cases over 7 months. The Procedural

TABLE 1 Zwisch scale of operative autonomy

	Attending behaviors	Resident behaviors	Autonomy scale level
Show and tell	<ul style="list-style-type: none"> Does majority of key portions as the surgeon Narrates the case, anatomy, skills 	<ul style="list-style-type: none"> Opens and closes First assists, observes 	1
Active help	<ul style="list-style-type: none"> Leads actively for >50% of critical portion of case Identifies key anatomy, optimizes the surgical field Coaches technical skills 	<ul style="list-style-type: none"> Actively assists Practice component skills 	2
Passive help	<ul style="list-style-type: none"> Follows resident's lead for >50% of critical portion Coaches for polish/refinement 	<ul style="list-style-type: none"> Recognizes transition points Accomplishes next steps 	3
Supervision only	<ul style="list-style-type: none"> Gives no unsolicited advice for >50% of critical portion Monitors progress and patient safety 	<ul style="list-style-type: none"> Mimics independence Recovers from most errors 	4

Source: Adapted from Kozin et al.²⁵

Learning and Safety Collaborative is a nonprofit organization that aims to improve the evaluation of operative experiences.¹¹ The culmination of their research has been the novel smartphone application SIMPL that allows residents to receive timely feedback for every case performed. Following an operative procedure, the trainee initiates the evaluation by documenting the date and the type of operation; the resident and faculty are then notified to complete same evaluation consisting of the three structured questions using the application's minimalist interface (Figure 1). Assessments expire in 72 hours as research suggests that later assessments do not accurately reflect performance.¹² Participants were asked to start their evaluations as soon as possible after each case to reduce recall bias. Faculty/residents cannot see the other party's evaluation until both have been submitted.

Three structured questions on SIMPL log a resident's operative autonomy achieved, their overall performance, and the case complexity. In this study, we focused on the question of operative autonomy, which was tracked on a four-level Zwisch scale. The Zwisch scale has been validated to reliably describe trainee surgical autonomy.^{13,14} The first level is "show and tell," during which the attending performs the majority of the critical portions of the case as the resident observes. Higher levels on the Zwisch scale suggest increasing autonomy

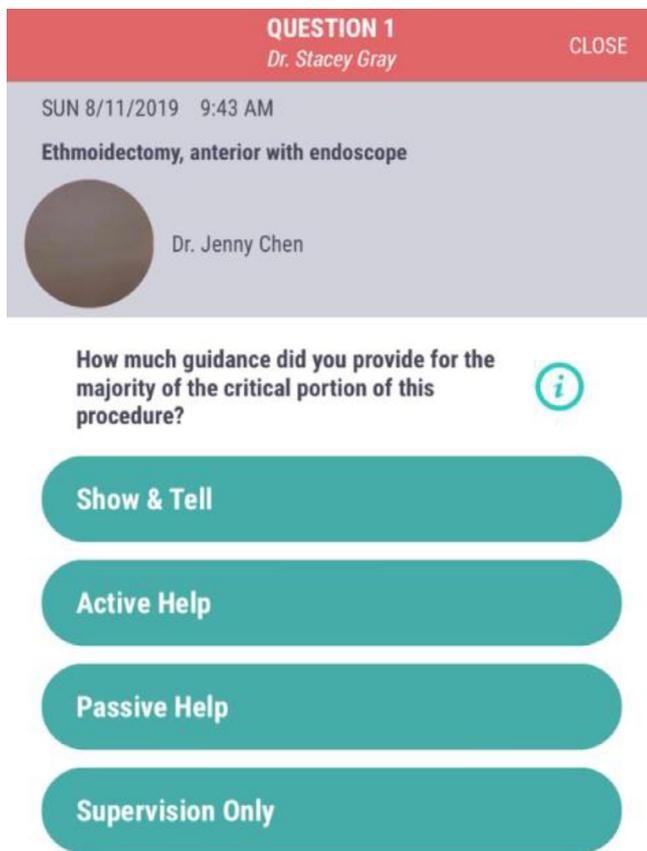


FIGURE 1 Screenshot of the SIMPL interface. This is the first out of three structured questions, which asks for a rating of resident autonomy on the Zwisch scale. Clicking on the italicized "i" icon brings the user to a table describing the different levels on the scale. SIMPL, system for improving and measuring procedural learning

("active help" and "passive help") and at the highest level of "supervision only," the resident safely simulates independent practice and the attending observes without offering significant input on surgical decision-making (Table 1). The second and third structured questions of SIMPL log a resident's performance on a 5-level scale (from "critical deficiency" to "exceptional performance"; Table 2) and case complexity on a 3-level scale (easiest 1/3, average 1/3, hardest 1/3 of cases for a given type of procedure). Note that for cases performed at the "show and tell level," no assessment of performance was logged. Software from the SIMPL application was modified from a General Surgery platform to focus on specific otolaryngology cases, as previously described.²⁵

2.3 | Statistical methods

Descriptive statistics were used to examine SIMPL and survey data. For survey data, Mann-Whitney *U* tests were used to compare levels of autonomy expected by residents and faculty for key indicator cases. For SIMPL, primary outcomes were the average Zwisch scale ratings reported by faculty and residents. For confidentiality purposes, analysis was not performed if there were either fewer than two residents at the same training level who performed the procedure or if there were fewer than five procedures performed by residents at the same training level. Paired comparisons were between resident and faculty ratings were made with Wilcoxon signed-rank test; where the comparison was significant, the rank correlation was calculated using the difference proportion of the two rank sums.¹⁵ Comparisons of

TABLE 2 Operative performance scale

	Performance descriptor	Performance scale level
Unprepared/critical deficiency	Resident is poorly prepared to perform this procedure	1
Inexperienced with procedure	Resident appears inexperienced in performing this procedure with frequent problems with technique, execution, forward planning	2
Intermediate performance	Resident is at an intermediate stage of development; performance of procedural elements is variable but acceptable for the amount of experience with procedure	3
Practice ready performance	Resident is ready to perform this 4 operation safely and independently assuming resident consistently performs procedure in this manner	4
Exceptional performance	Resident performs above the level expected of graduating residents	5

Source: Adapted from Kozin et al.²⁵

survey predicted autonomy and actual autonomy achieved logged on SIMPL were made with Mann-Whitney *U* tests. Correlations between resident autonomy and other factors (resident performance, resident level) were calculated with Spearman's rank correlation coefficient. Statistical analyses were performed using MedCalc (Ostend, Belgium) and Graphpad Prism (La Jolla, CA).

3 | RESULTS

3.1 | Study demographics

Twenty-three residents and 17 faculty members used the SIMPL smartphone application between 1st December 2017 and 30th June 2018. Residents spanning a range of training experience (PGY1-5) and faculty representing all major subspecialties of otolaryngology participated.

3.2 | Survey on resident autonomy

The response rate for the survey was 83% (82% attendings, 83% residents). For all ACGME selected key indicator cases, participants responded that increasing autonomy should be given with higher levels of training. For all cases, residents and faculty agreed that PGY5 residents should operate at a level between "passive help" and "supervision only." In the survey, residents overestimated the autonomy that should be achieved by PGY5 residents compared with faculty for 4 of 14 key indicator procedures: parotidectomy ($P = .012$), rhinoplasty ($P = .0053$), adult and pediatric airway procedures ($P = .0063$), and thyroid/parathyroidectomy ($P = .025$). For remaining key indicator case in the survey, residents and faculty predicted the autonomy of PGY5 residents to be at similar levels (Figure 2).

3.3 | SIMPL

Over the study period, 833 total evaluations were logged. Attending response rate to resident-initiated evaluations was 75% across all cases. All participants logged at least one case during the study. The median number of assessments logged per resident was 14 and the median number logged per attending was 20.¹⁶ Seventy-four percent of residents logged at least 10 cases, and 76% of faculty logged at least 10 cases.

3.4 | Case complexity

Case complexity was rated as "average" by participants in 74% ($n = 620$) of all cases logged. Fewer cases were rated as the easiest 1/3 of cases (8%, $n = 70$) or the hardest 1/3 of cases (17%, $n = 143$). There was no significant association between level of resident training and the complexity of cases rated by attendings (Spearman's correlation $P > .05$), suggesting that although senior residents were not disproportionately being exposed to more complex cases.

3.5 | Autonomy

Across all 833 evaluations, autonomy was associated with level of resident training (Spearman's correlation $r = 0.51$, $P < .0001$) and resident level of performance ($r = 0.47$, $P < .0001$). Autonomy was inversely associated with level of complexity ($r = -0.11$, $P = .0014$).

Excluding unilateral SIMPL evaluations, there were 356 paired evaluations, of which 253 were for ACGME selected key indicator cases (71%). The most common key indicator procedures were ethmoidectomy (24%, $n = 60$), rhinoplasty (15%, $n = 39$), parotidectomy (14%, $n = 35$), bronchoscopy (10%, $n = 25$), and mastoidectomy (6%,

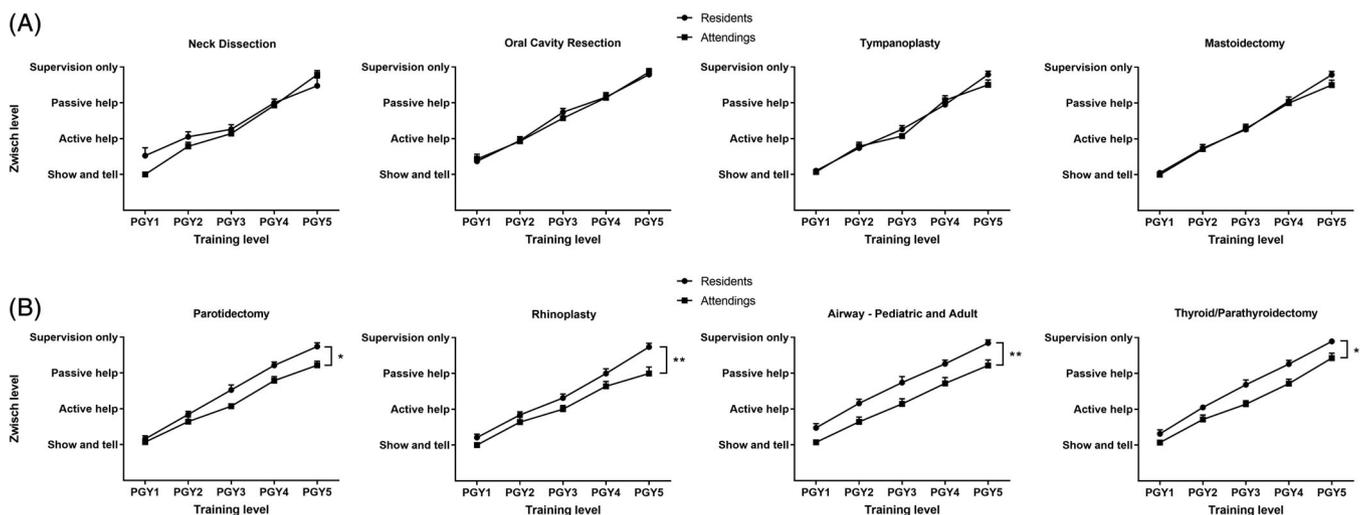


FIGURE 2 Expected levels of resident autonomy. A, For most key indicator cases, residents and faculty agreed on the level of autonomy expected. B, For a small number of key indicator cases (4 of 14), residents and faculty had different expectations on the level of autonomy achievable by PGY5. * $P < .05$, ** $P < .01$ for comparisons at PGY5. Error bars show standard error of the mean (SEM)

TABLE 3 Otolaryngology key indicator cases logged in SIMPL

Key indicator cases	Paired evaluations logged (% , total n = 253)
Ethmoidectomy	60 (24%)
Rhinoplasty	39 (15%)
Parotidectomy	35 (14%)
Bronchoscopy	25 (10%)
Mastoidectomy	15 (6%)
Neck dissection	13 (5%)
Oral cavity resection	12 (5%)
Flaps and grafts	11 (4%)
Airway	11 (4%)
Tympanoplasty	10 (4%)
Stapedectomy/ossiculoplasty	9 (4%)
Congenital neck masses	7 (3%)
Thyroidectomy/parathyroidectomy	4 (2%)
Mandible/maxilla (facial trauma)	2 (1%)

Abbreviation: SIMPL, system for improving and measuring procedural learning.

n = 15) (Table 3). Attendings' ratings of resident autonomy for rhinoplasty and ethmoidectomy were significantly higher compared with residents (Wilcoxon signed-rank test $P = .026$ and $P = .039$, respectively). The rank correlation of the two comparisons were $r = 0.65$ and $r = 0.46$, respectively, indicating higher ratings from attendings. There were no statistically significant differences in autonomy ratings logged in SIMPL by faculty compared with residents for other ACGME selected key indicator cases.

3.6 | Comparing survey predicted and actual autonomy

Comparing survey responses of predicted autonomy to actual autonomy logged on SIMPL, for most procedures (10 of 14, 71%), no statistically significant difference could be detected between residents or faculty's surveyed expectations and the level of autonomy achieved in the operating room. For two procedures (mastoidectomy and ethmoidectomy), residents and faculty both overpredicted the level of autonomy achieved by senior residents in the survey (Table 4, Figure 3A). For two

TABLE 4 Surveyed expectations for resident autonomy compared with ratings logged on SIMPL for key indicator cases (mean Zwisch levels)

Key indicator cases	Resident surveyed expectations of autonomy	Resident ratings of autonomy on SIMPL	Comparison	Attending surveyed expectations of autonomy	Attending ratings of resident autonomy on SIMPL	Comparison
Ethmoidectomy						
PGY2	2.1	2.0	$P > .05$	1.8	2.0	$P > .05$
PGY4	3.4	2.4	$P < .0001$	3.1	2.7	$P < .0001$
PGY5	3.8	3.0	$P = .0001$	3.6	3.1	$P = .024$
Rhinoplasty						
PGY2	1.8	1.3	$P > .05$	1.6	1.7	$P > .05$ $P > .05$
PGY5	3.7	3.0	$P = .017$	3.0	3.0	
Parotidectomy						
PGY3	2.5	2.0	$P = .044$	2.0	2.3	$P > .05$
PGY4	3.2	2.5	$P = .009$	2.8	2.4	$P > .05$
PGY5	3.7	2.7	$P = .007$	3.2	3.0	$P > .05$
Bronchoscopy						
PGY2	2.5	2.9	$P > .05$	1.9	2.7	$P > .05$
Mastoidectomy						
PGY5	3.8	3.1	$P = .002$	3.5	2.9	$P = .029$
Neck dissection						
PGY3	2.3	1.6	$P > .05$	2.1	2.2	$P > .05$
Flaps and grafts						
PGY2	1.9	1.7	$P > .05$	1.8	1.7	$P > .05$
Airway						
PGY2	2.2	1.8	$P > .05$	1.6	2.0	$P > .05$
Stapedectomy/ossiculoplasty						
PGY5	3.6	3.1	$P > .05$	3.3	2.4	$P > .05$

Abbreviation: SIMPL, system for improving and measuring procedural learning.

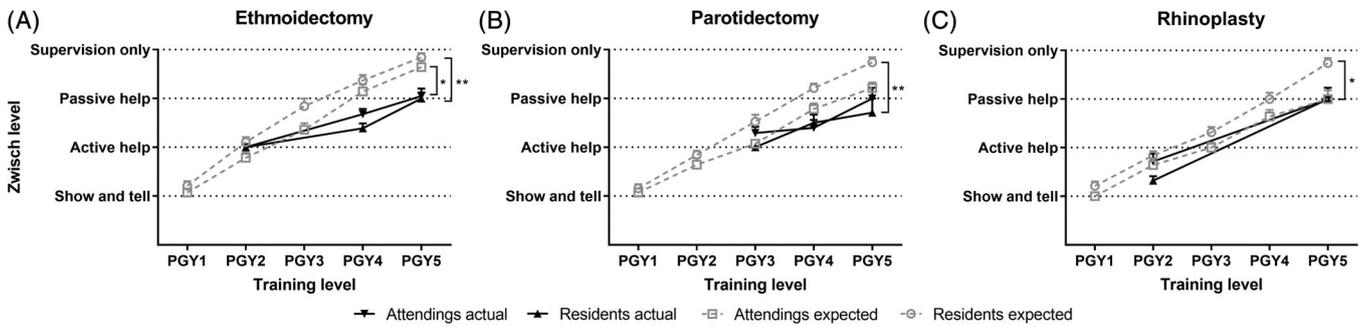


FIGURE 3 Comparing expected and actual levels of autonomy. A, For a small number of procedures such as ethmoidectomy, residents and faculty all over predicted the level of autonomy achieved by senior residents on surveys. B and C, For parotidectomy and rhinoplasty, residents overestimated the level of autonomy achieved by senior residents while attendings' surveyed expectations were more realistic. * $P < .05$, ** $P < .01$ for comparisons at PGY5. Error bars show SEM

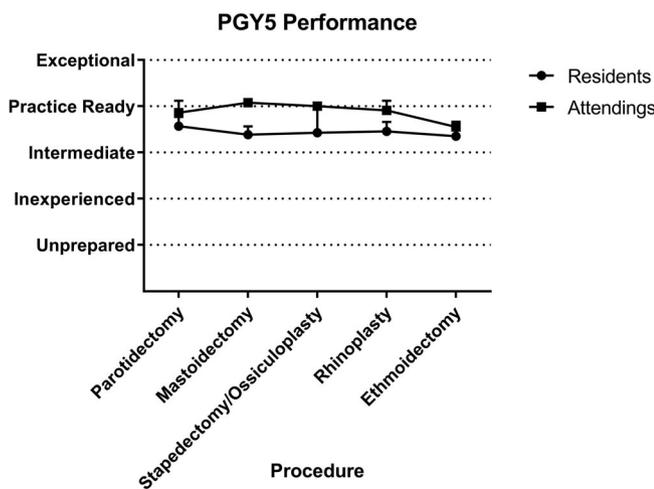


FIGURE 4 Performance of senior residents. Faculty felt that PGY5 residents performed cases well and were practice ready or nearly practice ready for all cases where this analysis could be conducted. Error bars show SEM

procedures (parotidectomy and rhinoplasty), residents overpredicted the level of autonomy senior residents would achieve in the survey, while attendings' predictions were similar to actual data logged on SIMPL (Table 4, Figure 3B,C).

3.7 | Performance

Faculty felt that PGY5 residents performed cases well and were practice ready or nearly practice ready for all cases where this analysis could be conducted (Figure 4). Residents appeared to rate themselves at slightly lower levels of performance but this comparison was significant only for mastoidectomy (Wilcoxon signed-rank test, $P < .05$).

4 | DISCUSSION

This is the first prospective study in otolaryngology to compare survey predicted and actual levels of resident surgical autonomy logged using

a mobile-based application. In this single-center pilot study, we found that residents and faculty reported different expectations of autonomy for a number of key indicator cases in surveys, for which residents are prone to overestimate the level of independence achieved by their last year of training. Comparing predicted autonomy as reported in surveys with actual independence achieved in the operating room logged on SIMPL, residents overestimated their autonomy as senior residents for at least four of 14 key indicator cases while attendings overestimated this for at least two of these same procedures.

Although it would be ideal for residents to consistently simulate independent practice before graduation, this study shows that this might not always be achieved for complex otolaryngology surgeries. For some procedures, such as parotidectomy and rhinoplasty, residents and faculty disagreed on the level of autonomy that should be allotted to trainees, which might be due to the fact that these surgeries have higher risk complications or cosmetic implications. In these cases, it is possible that barriers to autonomy exist because of little room for surgical error.

For other procedures, an autonomy gap exists despite faculty and residents' agreement on surveys that senior residents should operate with supervision only, such as for ethmoidectomy and mastoidectomy. Further inquiry into why this autonomy gap exists is necessary to determine if it is the expectations or the actual practice that should be addressed. One factor this current study could not adequately control was the natural stratification of cases by resident level that leaves senior residents with a disproportionate number of difficult cases. Although there was no significant correlation between the complexity of cases assigned and postgraduate year level, this would be difficult to discern because attendings rated the majority of cases as "average complexity" for the given type of procedure. Additionally, it is interesting to note that, although both ethmoidectomy and mastoidectomy procedures utilize microscopes or endoscopes that allow attendings to watch residents perform the cases, the field of vision is limited and attendings have no tactile feedback, making such cases particularly difficult to teach.¹⁷ Therefore, perhaps these procedures are inherently less likely to be "supervision only" as attendings may feel the need to periodically take over the instruments and inspect residents'

work. In this study, cases that use endoscopes or microscopes accounted for 31% (n = 256) of total assessments including 19% (n = 159) that were ethmoidectomies or mastoidectomies. Future survey or qualitative studies should investigate why these types of procedures are less likely to result in high resident operative autonomy.

Outside of these explanations, there are a variety of case-specific factors that affect how residents and faculty negotiate autonomy in the operating room. In this study, performance and training level were correlated with autonomy but other factors such as supervising surgeons' experience teaching specific procedures, prior operative interactions between attendings and residents, as well as residents' personal attributes are also potential contributing factors.¹⁸ Case details such as patient comorbidities or patient preference may also limit resident participation. Resident participation may also increase operative time and hospital costs,³ and attending surgeons may also feel varying degrees of personal responsibility to their patients who expect their surgeons to participate in their procedure.¹⁹

Although patient safety is of paramount importance, it is desirable to maximize residents' operative experiences to train the next generation of surgeons. This current study begins to characterize the autonomy gap in otolaryngology and offers vocabulary through the Zwisch scale with which residents and faculty can negotiate participation in the operating room. The act of logging resident experiences with an emphasis on tracking autonomy and performance may cause faculty and residents to be more mindful of educational opportunities in the operating room. In future studies, a more detailed understanding of the factors that unduly limit resident learning may further improve the quality of surgical training.

It is beyond the scope of this study to determine whether the level of autonomy senior residents logged correlates with readiness for independent practice. Notably, although residents in this study did not always operate with "supervision only" by the end of training, the performance ratings given to PGY5 residents are reassuring, as faculty found them consistently "practice ready." Although the current study sample was powered to study only a handful of ACGME selected key indicator cases, the data suggest residents perform a variety of otolaryngology procedures well by the end of training. The further question of how autonomy as a resident relates to performance as a junior attending will be critical. Specific information about graduates' abilities to confidently perform general otolaryngology procedures during early years of independent practice has not been examined. Tracking experiences and outcomes in early years of independent practice may be necessary to understand the impact of autonomy during training.

While this is the first study in otolaryngology to track resident surgical autonomy using a mobile application, its findings are similar to those in general surgery. A study of general surgery residents at a single residency program found that autonomy logged using SIMPL was significantly less than surveyed expectations for chief residents for many core procedures (eg, laparoscopic appendectomy/cholecystectomy/partial colectomy, hemorrhoidectomy, open inguinal hernia repair).²⁰ A multicenter trial of 14 general surgery programs describing the progression of trainee operative autonomy during training found that residents in their final 6 months of training reached meaningful autonomy in "core procedures" defined by the American Board of

Surgery in 77.4% of cases.²¹ These authors concluded that residents are not universally ready to independently perform some common procedures by graduation. However, it is currently unclear how much autonomy is actually necessary for residents to achieve competency, and how many observations are needed to make generalizable assessments about the degree of autonomy attained.^{21,22} A follow-up study examining cases logged using SIMPL at 14 general surgery programs found that the factors that correlated with autonomy were (in order of the strength of the predictor) quality of a resident's operative performance during the case, typical autonomy granted by the supervising surgeon, case difficulty, and trainee experience level.²³

4.1 | Study limitations

The primary limitation of this study is the single institution design that reflects the opinions and experiences of a limited number of residents and faculty. We were unable to obtain the quantity of evaluations needed to perform analysis on all the types of procedures that otolaryngology residents perform beyond common key indicator cases. Moreover, as a single center study, generalizations to other residency programs may or may not be valid as the case mix, surgical staff, and teaching ethos may differ at programs across the country.

The limited sample size in this study made it difficult to track and analyze autonomy at all different levels of residency for all key indicator surgeries. For example, at the study institution, key indicator otologic cases are performed primarily by PGY4 and five residents on their otology rotations. Small numbers of certain procedures logged (eg, facial trauma cases, thyroid/parathyroidectomy, and congenital neck masses) precluded statistical analyses for those cases. Due to the limited sample size, we are also not able to stratify the data to analyze other questions such as whether PGY5 residents are able to achieve autonomy at the end of their rotations or at the very end of residency. Stratifying such data from future multicenter trials may also reveal how a multitude of other factors affects surgical autonomy.

Additionally, the data logged could be confounded by recall biases. It can be difficult for subjects to separate individual components of complex cases from the larger surgery performed; for example, a senior resident who performs an ethmoidectomy during an operation with a difficult frontal sinus dissection may evaluate their autonomy and performance based on the memory of the more difficult portion of the case rather than solely on the ethmoidectomy. This is a limitation inherent to research dependent on participants' evaluations of each other rather than on objective third party assessments.

Furthermore, although inter-rater reliability has been studied for SIMPL in the field of general surgery,⁹ this has not been assessed for otolaryngology faculty. Future research will determine if similar findings of inter-rater reliability are noted in otolaryngology. Videos used for frame-of-reference training at the start of this study also used general surgery procedures; future studies should develop training materials specific to otolaryngology.

In the survey study, a possible limitation is question fatigue causing respondents to potentially respond to each procedure with the

same level of autonomy for a given PGY level. However, looking at individual survey responses, only one participant out of 17 gave “straight-lined” answers for each PGY year while the remaining subjects varied responses for different procedures at each training level.

Last, this study has not explored the ways in which SIMPL could be used to complement other tools useful for the assessment of surgical trainees. These tools include Objective Structured Assessment of Technical Skills (OSATS) and other modalities of evaluations including ACGME Milestones surveys,²⁴ MedHub, or New Innovations. Ideally, programs would use SIMPL to inform other evaluations and allow programs to deliver both granular and general feedback to residents. Future studies using alternative assessment tools and standardized independent observers may be valuable in providing validity evidence to support the use of SIMPL.

5 | CONCLUSIONS

This single-center pilot study finds that otolaryngology residents and faculty do not always agree on the level of surgical autonomy that residents should expect to achieve during training. Additionally, the degree of actual surgical autonomy granted to trainees varies depending on the procedure performed. Multicenter studies will be critical to explore the factors that impact operative autonomy during residency, understand the degree of autonomy necessary to ensure competency by graduation, and determine if simulated autonomy during training translates into improved outcomes during the early years of independent practice.

DISCLOSURES

Dr Brian George and Dr Jordan Bohnen are unpaid, volunteer members of the Procedural Learning and Safety Collaborative, a nonprofit surgical education consortium based in Boston, MA. This work was completed in collaboration with the Procedural Learning and Safety Collaborative.

ACKNOWLEDGMENTS

This research would not have been possible without the support of the faculty and residents at our institution. Specifically, we are indebted to B. Bleier, T. Hadlock, C. Hartnick, E. Holbrook, D. Jung, D. Lee, L. Lee, R. Lindsay, A. Quesnel, J. Richmon, A. Sedaghat, D. Shaye, P. Song, M. Varvares, and D. B. Welling. We are deeply indebted to every resident at the Harvard Otolaryngology Residency Program.

ORCID

Jenny X. Chen  <https://orcid.org/0000-0002-2992-4771>

Elliott Kozin  <https://orcid.org/0000-0002-0305-0682>

REFERENCES

- Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system — rationale and benefits. *N Engl J Med*. 2012;366(11):1051-1056. <https://doi.org/10.1056/NEJMSr1200117>.
- Brunworth JD, Sindwani R. Impact of duty hour restrictions on otolaryngology training: divergent resident and faculty perspectives. *Laryngoscope*. 2006;116(7):1127-1130. <https://doi.org/10.1097/01.mlg.0000224348.44616.fb>.
- Puram SV, Kozin ED, Sethi R, et al. Impact of resident surgeons on procedure length based on common pediatric otolaryngology cases. *Laryngoscope*. 2015;125(4):991-997. <https://doi.org/10.1002/lary.24912>.
- Puram SV, Barber SR, Kozin ED, et al. Outcomes following pediatric auditory brainstem implant surgery: early experiences in a North American center. *Otolaryngol Head Neck Surg*. 2016;155(1):133-138. <https://doi.org/10.1177/0194599816637599>.
- Kauvar DS, Braswell A, Brown BD, Harnisch M. Influence of resident and attending surgeon seniority on operative performance in laparoscopic Cholecystectomy. *J Surg Res*. 2006;132(2):159-163. <https://doi.org/10.1016/j.jss.2005.11.578>.
- Bartels DD, McCann ME, Davidson AJ, Polaner DM, Whitlock EL, Bateman BT. Estimating pediatric general anesthesia exposure: quantifying duration and risk. *Paediatr Anaesth*. 2018;28(6):520-527. <https://doi.org/10.1111/pan.13391>.
- Dine CJ, Myers JS. Balancing Supervision and Autonomy: An Ongoing Tension. AHRQ Patient Safety Network. <https://psnet.ahrq.gov/perspectives/perspective/116/balancing-supervision-and-autonomy-an-ongoing-tension>. Accessed July 14, 2018.
- Brown DJ, Thompson RE, Bhatti NI. Assessment of operative competency in otolaryngology residency: survey of US Program Directors. *Laryngoscope*. 2008;118(10):1761-1764. <https://doi.org/10.1097/MLG.0b013e31817e2c62>.
- George BC, Teitelbaum EN, Darosa DA, et al. Duration of faculty training needed to ensure reliable or performance ratings. *J Surg Educ*. 2013;70(6):703-708. <https://doi.org/10.1016/j.jsurg.2013.06.015>.
- Review Committee for Otolaryngology - Head and Neck Surgery. Required Minimum Number of Key Indicator Procedures for Graduating Residents. 2018. http://www.acgme.org/Portals/0/PFAssets/ProgramResources/280_Required_Minimum_Number_of_Key_Indicator_Procedures.pdf. Accessed August 11, 2019.
- Home - PLSC. <http://www.procedurallearning.org/>. Accessed January 1, 2017.
- Williams RG, Chen XP, Sanfey H, Markwell SJ, Mellinger JD, Dunnington GL. The measured effect of delay in completing operative performance ratings on clarity and detail of ratings assigned. *J Surg Educ*. 2014;71(6):e132-e138. <https://doi.org/10.1016/j.jsurg.2014.06.015>.
- DaRosa DA, Zwischenberger JB, Meyerson SL, et al. A theory-based model for teaching and assessing residents in the operating room. *J Surg Educ*. 2013;70(1):24-30. <https://doi.org/10.1016/j.jsurg.2012.07.007>.
- George BC, Teitelbaum EN, Meyerson SL, et al. Reliability, validity, and feasibility of the Zwisch scale for the assessment of intraoperative performance. *J Surg Educ*. 2014;71(6):e90-e96. <https://doi.org/10.1016/j.jsurg.2014.06.018>.
- Kerby DS. The simple difference formula: an approach to teaching nonparametric correlation. *Compr Psychol*. 2014;3:11.IT.3.1. <https://doi.org/10.2466/11.IT.3.1>.
- Chen JX, Kozin ED, Bohnen JD, et al. Assessments of otolaryngology resident operative experiences using mobile technology: a pilot study. *Otolaryngol Head Neck Surg*. 2019; In Press. <https://doi.org/10.1177/0194599819868165>.
- Gross RD, Sheridan MF, Burgess LP. Endoscopic sinus surgery complications in residency. *Laryngoscope*. 1997;107(8):1080-1085. <https://doi.org/10.1097/00005537-199708000-00014>.
- Meyerson SL, Sternbach JM, Zwischenberger JB, Bender EM. The effect of gender on resident autonomy in the operating room. *J Surg Educ*. 2017;74(6):e111-e118. <https://doi.org/10.1016/j.jsurg.2017.06.014>.

19. Moore FD. Ethical problems special to surgery: surgical teaching, surgical innovation, and the surgeon in managed care. *Arch Surg*. 2000;135(1):14-16. <https://doi.org/10.1001/archsurg.135.1.14>.
20. Meyerson SL, Teitelbaum EN, George BC, Schuller MC, DaRosa DA, Fryer JP. Defining the autonomy gap: when expectations do not meet reality in the operating room. *J Surg Educ*. 2014;71(6):e64-e72. <https://doi.org/10.1016/j.jsurg.2014.05.002>.
21. George BC, Bohnen JD, Williams RG, et al. Readiness of US general surgery residents for independent practice. *Ann Surg*. 2017;266(4):582-594. <https://doi.org/10.1097/SLA.0000000000002414>.
22. Williams RG, Swanson DB, Fryer JP, et al. How many observations are needed to assess a surgical trainee's state of operative competency? *Ann Surg*. 2019;269(2):377-382. <https://doi.org/10.1097/SLA.0000000000002554>.
23. Williams RG, George BC, Meyerson SL, et al. What factors influence attending surgeon decisions about resident autonomy in the operating room? *Surgery*. 2017;162(6):1314-1319. <https://doi.org/10.1016/j.surg.2017.07.028>.
24. Tsue TT. Developing the otolaryngology milestones. *J Grad Med Educ*. 2014;6(1 suppl 1):162-165. <https://doi.org/10.4300/JGME-06-01s1-21>.
25. Kozin ED, Bohnen JD, George BC, et al. Novel mobile app allows for fast and validated intraoperative assessment of otolaryngology residents. *OTO Open*. 2017;1(1):2473974X16685705. <https://doi.org/10.1177/2473974X16685705>.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Chen JX, Kozin E, Bohnen J, et al. Tracking operative autonomy and performance in otolaryngology training using smartphone technology: A single institution pilot study. *Laryngoscope Investigative Otolaryngology*. 2019;4:578-586. <https://doi.org/10.1002/lio2.323>