Looking Beyond the Numbers: A Comparison of Operative Self-Efficacy, Supervision, and Case Volume in General Surgery Residency

Rachel M. Jensen, MD, MHPE
Ananya Anand, MD
LaDonna E. Kearse, MD
James R. Korndorffer Jr, MD, MHPE

ABSTRACT

Background A national survey of general surgery residents revealed significant self-assessed deficits in preparation for independent practice, with only 7.7% of graduating postgraduate year 5 residents (n=1145) reporting self-efficacy for all 10 commonly performed operations surveyed.

Objective We sought to understand why this phenomenon occurs. We hypothesized that self-efficacy would be positively correlated with both operative independence and case volume.

Methods We compared 3 independent datasets: case information for the same 10 previously surveyed operations for residents graduating in 2020 (dataset 1), operative independence data obtained through the SIMPL OR app, an operative self-assessment tool (dataset 2), and case volume data obtained through the Accreditation Council for Graduate Medical Education National Data Report (dataset 3). Operations were categorized into high, middle (mid), and low self-efficacy tiers; analysis of variance was used to compare operative independence and case volume per tier.

Results There were significant differences in self-efficacy between high (87.7%), mid (68.3%), and low (25.4%) tiers (P=.008 [95% CI 6.2, 32.7] for high vs mid, P<.001 for high vs low [49.1, 75.6], and P<.001 for mid vs low [28.7, 57.1]). The percentage of cases completed with operative independence followed similar trends (high 32.7%, mid 13.8%, low 4.9%, P=.006 [6.4, 31.4] for high vs mid, P<.001 [15.3, 40.3] for high vs low, P=.23 [-4.5, 22.3] for mid vs low). The total volume of cases decreased from high to mid to low self-efficacy tiers (average 91.8 to 20.8 to 11.1) but did not reach statistical significance on post-hoc analysis.

Conclusions In this analysis of US surgical residents, operative independence was strongly correlated with self-efficacy.

Introduction

The development of self-efficacy is essential for continued growth and lifelong learning, particularly in task- or procedurally oriented specialties such as surgery. Self-efficacy, or one's belief in their ability to accomplish a specific task, is a concept described by cognitive psychologist Albert Bandura in the 1970s and comes from a combination of mastery experiences, social persuasion (coaching, feedback), vicarious experiences, and psychological responses.[1] However, there is evidence to suggest that surgery residents may be lacking in operative self-efficacy. A survey on self-efficacy of graduating postgraduate year (PGY) 5 residents was linked to the 2020 American Board of Surgery In-Training Examination (ABSITE), to ensure that all PGY-5 residents who took the examination had the option to complete the survey. It revealed that only 7.7% of graduating residents reported self-efficacy in all 10 commonly performed operations.[2] This presents a challenge when considering graduating residents' readiness for independent clinical practice.

While this study identified a significant deficit in resident self-efficacy, it also presents further questions as to what is contributing to the deficit and how it can be addressed. Kearse et al attempted to answer this question and identified a combination of institutional and individual factors, such as male sex, emphasis on autonomy, and peer/faculty socialization associated with higher self-efficacy.[3] Performing operations with greater levels of resident responsibility and in a teaching assistant role were also correlated with higher self-efficacy.[3] While this takes a step toward better understanding and thereby addressing the self-efficacy deficit, it provides little information about case-level experiences that might shape a resident's perceptions of their own abilities. These case-level experiences include elements of both quantity and quality, such as number of cases performed, frequency of performance, and degree of operative independence when performing the case. At present, surgery residents log their operative cases as the

DOI: http://dx.doi.org/10.4300/JGME-D-23-00461.1

primary method to track operative experience during training. While such logs document case volume, they do not capture other elements of the case-level experiences that may be important to consider in preparing residents for independent clinical practice.

These case-level experiences are important to evaluate in the context of Bandura's self-efficacy theory, where there are many different facets that contribute to the development of self-efficacy.[1] Most of these facets are centered around prior experience, such number of times a case was performed, operative entrustment, vicarious experiences watching someone else perform an operation, and the physiological response associated with these experiences. Thus, the theory of self-efficacy would suggest that experience shapes perceptions. We sought to evaluate how operative independence and case volume contribute to self-efficacy; operative independence speaks to the quality of the experience, while case volume speaks more to quantity. We hypothesized that self-efficacy would be positively correlated with both operative independence and case volume.

KEY POINTS

What Is Known
Graduates of surgery residency programs report lower-than-ideal self-efficacy for 10 commonly performed operations, but it is not clear if that perception is related to other data assessing operative independence.

What Is New
This study compared data from 3 sources: self-reported self-efficacy, operative independence as assessed by residents and supervisors, and case volume. Self-efficacy was strongly correlated with operative independence.

Bottom Line
In order to increase residents' self-efficacy, attention must be paid to the quality of the operative experience as opposed to just case volume.

Methods

To make comparisons between operative self-efficacy, operative independence, and case volume, data were compared from 3 different and independent data sources as follows: (1) self-reported operative self-efficacy data obtained through a survey taken by graduating PGY-5 residents; (2) self-reported operative independence or autonomy data obtained through the SIMPL OR phone application (Procedural Learning and Safety Collaborative); and (3) case numbers obtained through the publicly available 2019-2020 Accreditation Council for Graduate Medical Education (ACGME) case log data. Throughout this study, *operation* refers to a specific surgery (eg, laparoscopic appendectomy or trauma thoracotomy), while *case* refers to a particular instance of performing an operation. To provide consistency between the distinctly unique datasets, data were obtained from general surgery residents graduating in 2020 for all 3 sources. We will now describe the 3 datasets in more detail.

Dataset 1

Self-efficacy survey data from the graduating general surgery PGY-5 cohort was obtained from an institutional review board (IRB)-approved survey included as a posttest section on the 2020 ABSITE. The ABSITE is offered annually, and survey questions are often included at the end of the examination; our institution spearheaded the self-efficacy survey in 2020. The authors selected a total of 10 common general surgery operations for the focus of the study, which included 4 laparoscopic cases (laparoscopic appendectomy, cholecystectomy, right hemicolectomy, and diagnostic laparoscopy) and 6 open cases (wide local excision, breast biopsy, thyroidectomy, inguinal hernia repair, trauma exploratory laparotomy, and trauma thoracotomy). Individuals were asked to rate their ability to perform these procedures without attending assistance at this time using a 5-point scale (1=not able to; 2=likely able to; 3=more likely able to; 4=most likely able to; 5=definitely able to). Consistent with previous literature, self-efficacy was considered to be a binary construct where a score of 5 was considered self-efficacious and a score of 1 to 4 was considered not self-efficacious. Complete methods for case selection and survey design are described in Anderson et al.[2]

Dataset 2

Data were also obtained through the SIMPL OR application on the same operations queried in the 2020 ABSITE survey through a research-based deidentified data use agreement. On the SIMPL OR application, residents report their perceptions of operative independence for individual *cases* (eg, a specific laparoscopic appendectomy). Thus, each resident could theoretically submit multiple individual cases for a given operation (eg, 3 trauma thoracotomies in 5 years). SIMPL OR is not used universally across all general surgery residency programs; however, the programs that do use SIMPL OR are diverse in program type (academic vs academic-affiliated vs independent). When using the SIMPL OR application, both trainees and attending surgeons answered 3 key questions related to the case regarding complexity, performance, and supervision. For supervision, individuals selected a score from 1 to 4 based on the Zwisch scale: 1=show and tell;

2=active help; 3=passive help; and 4=supervision only.[4] Case supervision data obtained from the 5 years of residency for 2020 graduates were reviewed for the aforementioned 10 operations because many of the operations included in the survey are more commonly performed at the junior resident stage; thus, operative independence could theoretically be achieved at any year of training. Operative independence was also considered to be a binary construct where a Zwisch score of 4 (supervision only) was considered independent and a Zwisch score of 1 to 3 was considered nonindependent.

Dataset 3

The publicly available ACGME 2019-2020 National Data Report was used to compare graduating resident mean case numbers (surgeon total including surgeon chief and surgeon junior) for the same 10 operations.[5] Case numbers reflect the average number of times (cases) a given operation was performed by a resident over 5 years of training. Cases logged through the ACGME Case Log System were not granular enough to be able to identify 3 of the 10 operations: diagnostic laparoscopy, wide local excision, and laparoscopic right hemicolectomy.

The 10 operations were grouped into 3 tiers (high, middle, and low) based on resident-reported self-efficacy. Descriptive statistics were used to evaluate the frequency of reported operative independence per operation and mean frequency of operative independence per self-efficacy tier. Descriptive statistics were also used to calculate mean case volume per self-efficacy tier. Analysis of variance (ANOVA) was used to assess for differences in the percentage of cases completed with operative independence and the number of cases completed per the ACGME National Data Report based on self-efficacy tier. The Bonferroni calculation was used for post-hoc analysis. Correlations were used to evaluate the relationships between self-efficacy, operative independence, and case volume. Independent samples proportions tests were used to compare trainee and attending perceptions of operative independence. All statistical analyses were conducted using SPSS statistical software (IBM Corp).

The study was reviewed and approved by the Stanford University IRB.

Results

The 10 operations previously described by Anderson et al were organized into 3 self-efficacy tiers (3 to 4 operations per tier) based on the data obtained through the post-ABSITE 2020 resident survey.[2] High self-efficacy operations included laparoscopic cholecystectomy, laparoscopic appendectomy, diagnostic laparoscopy, and wide local excision. The middle self-efficacy operations included breast biopsy, inguinal hernia, and trauma exploratory laparotomy. The low self-efficacy operations included thyroidectomy, laparoscopic right hemicolectomy, and trauma thoracotomy. See TABLE 1 for a breakdown of self-efficacy

TABLE 1
Self-Efficacy, Supervision, and Case Volume by Operation

Self-Efficacy Tier	Operation	% (n of N) of PGY-5s Reporting Self-Efficacy	% (n of N) of Cases Trainees Reported Completing With Operative Independence	Mean No. of Cases Completed
High	Laparoscopic cholecystectomy	84.1 (1001 of 1190)	32.7 (577 of 1763)	118.6
	Laparoscopic appendectomy	89.8 (1068 of 1189)	40.2 (293 of 728)	64.9
	Diagnostic laparoscopy	86.8 (1030 of 1187)	27 (30 of 111)	–
	Wide local excision	90.2 (1072 of 1188)	28 (40 of 143)	–
	Mean	87.7	32	91.8
Middle	Breast biopsy	77 (913 of 1186)	7.9 (5 of 63)	11.9
	Inguinal hernia	67.1 (799 of 1190)	17.5 (74 of 424)	42.8
	Trauma exploratory laparotomy	60.7 (722 of 1190)	15.2 (7 of 46)	7.8
	Mean	68.3	13.5	20.8
Low	Thyroidectomy	19.6 (233 of 1190)	2.6 (3 of 114)	19.3
	Laparoscopic right hemicolectomy	26.5 (315 of 1187)	2.8 (3 of 106)	–
	Trauma thoracotomy	30 (357 of 1190)	10 (1 of 10)	2.8
	Mean	25.4	5.1	11.1

Abbreviation: PGY, postgraduate year.

TABLE 2
Pearson Correlations Between Self-Efficacy, Case Supervision, and Case Volume

Variable	Self-Efficacy, r	Case Supervision, r	Case Volume, r
Self-efficacy	1.00	0.84[a]	0.61
Case supervision		1.00	0.79[b]
Case volume			1.00

[a] Significant at the 0.01 level; *P*=.002 [0.5, 1.0].
[b] Significant at the 0.05 level; *P*=.04 [0.1, 1.0].

per operation with sample sizes in addition to tiered means.

The percentage of cases which *residents* reported as completing with operative independence was calculated based on SIMPL OR data for the same 10 operations.[4] The sample sizes vary by operation; see TABLE 1 for a full breakdown of operative independence per operation including sample sizes.

The mean number of cases reported per graduating resident (N=1257) on the 2020 ACGME National Data Report was reviewed for 7 of 10 operations. Due to lack of granularity in case reporting, diagnostic laparoscopy, wide local excision, and laparoscopic right hemicolectomy were excluded from the mean calculations. See TABLE 1 for a full breakdown of case volume.

ANOVA revealed significant differences in mean self-efficacy between all tiers: high vs mid (87.7% vs 68.3%, *P*=.008 [95% CI 6.2, 32.7]), high vs low (87.7% vs 25.4%, *P*<.001 [49.1, 75.6]), and mid vs low (68.3% vs 25.4%, *P*<.001 [28.7, 57.1]). The percentage of cases completed with operative independence followed similar trends with 32.7% vs 13.8%, *P*=.006 [6.4, 31.4] for high vs mid; 32.7% vs 4.9%, *P*<.001 [15.3, 40.3] for high vs low; and 13.8% vs 4.9%, *P*=.23 [-4.5, 22.3] for mid vs low. While the total volume of cases on average decreased from high (91.8) to mid (20.8) to low (11.1) self-efficacy tiers, this did not reach statistical significance on post-hoc analysis (*P*=.10 [-16.0, 157.8] for high vs mid, *P*=.09 [-14.4, 175.9] for high vs low, *P*>.99 [-77.1, 96.7] for mid vs low).

Correlations were calculated to evaluate the relationships between self-efficacy, supervision, and case volume. Self-efficacy and supervision showed a strong correlation (r=0.84, *P*=.002 [0.5, 1.0]). Self-efficacy and case volume showed no significant correlation (r=0.61, *P*=.14 [-0.3, 0.9]). Supervision and case volume showed a strong significant correlation (r=0.79, *P*=.04 [0.1, 1.0]). See TABLE 2 for a summary of these correlations.

Attending surgeon perceptions of operative independence demonstrated similar trends by self-efficacy tier as compared to resident perceptions: while attending surgeons generally perceived higher degrees of operative independence, this was statistically significant for only 3 of 10 procedures (laparoscopic cholecystectomy, laparoscopic appendectomy, and thyroidectomy); see TABLE 3. The Pearson correlation between resident and attending perceptions of operative independence was strong (r=0.95, *P*<.001 [0.8, 1.0]). The Pearson correlation between attending-reported operative independence and resident-reported self-efficacy was also strong (r=0.82, *P*=.003 [0.4, 1.0]).

Discussion

These study findings support the idea that prior case-specific experience impacts perceptions of self-efficacy. The strong correlation between level of supervision and self-efficacy is logical given the role of social persuasion (coaching, feedback, entrustment, etc) as an important contributor in self-efficacy theory. While there are certainly many different factors that contribute to a resident's perception of their self-efficacy, these findings suggest that level of supervision is one of many potential targets for improving the operative self-efficacy deficit. When residents have more opportunities for operative independence, they can further develop their own self-efficacy, which is essential for personal growth. This creates a positive self-efficacy cycle through social persuasion with further opportunities for operative independence as depicted in the FIGURE.

Though the study findings support the self-efficacy/operative independence relationship, it is worth noting the stark differences in the percentages of residents reporting procedural self-efficacy compared to the percentage of cases being completed with operative independence. A Zwisch score of 4 likely underestimates true self-efficacy, as there could be several reasons why a self-efficacious resident does not perform a case with operative independence (faculty preference, patient preference, time constraints, billing requirements, institution requirements, etc). Additionally, cases were included from all years of training, which should theoretically represent a range and continuum of independence levels. Despite this, these differences do suggest that some self-efficacious residents are not being given opportunities to perform cases

TABLE 3
Trainee vs Attending Perceptions of Operative Independence

Self-Efficacy Tier	Operation	% (n of N) of Cases Trainees Reported as Completing With Operative Independence	% (n of N) of Cases Attendings Reported as Completing With Operative Independence	*P* value (95% CI)
High	Laparoscopic cholecystectomy	32.7 (577 of 1763)	39.3 (534 of 1360)	<.001 (0.03, 1.0)
	Laparoscopic appendectomy	40.2 (293 of 728)	47.8 (288 of 602)	.01 (0.02, 0.13)
	Diagnostic laparoscopy	27 (30 of 111)	37.4 (40 of 107)	.10 (-0.21, 0.22)
	Wide local excision	28 (40 of 143)	35.8 (39 of 109)	.19 (-0.04, 0.19)
Middle	Breast biopsy	7.9 (5 of 63)	17.8 (8 of 45)	.12 (-0.03, 0.23)
	Inguinal hernia	17.5 (74 of 424)	22.8 (74 of 324)	.07 (-0.004, 0.11)
	Trauma exploratory laparotomy	15.2 (7 of 46)	31.7 (13 of 41)	.07 (-0.02, 0.33)
Low	Thyroidectomy	2.6 (3 of 114)	16.9 (12 of 71)	<.001 (0.05, 0.24)
	Laparoscopic right hemicolectomy	2.8 (3 of 106)	7.1 (7 of 98)	.15 (-0.02, 0.11)
	Trauma thoracotomy	10 (1 of 10)	14.3 (1 of 7)	.79 (-0.29, 0.40)

with operative independence, thereby inhibiting further personal growth. It also suggests that there are likely factors outside of operative independence that contribute to development of self-efficacy.

Case volume is one such factor that may affect self-efficacy through mastery experiences, despite not showing a significant correlation with self-efficacy in this study. It did, however, show a strong correlation with the level of supervision, suggesting that the number of cases completed likely contributes to the amount of operative independence granted. It is difficult to ascertain whether this is due to competence, performance, level of training, resident-faculty relationship, or other trainee- or institution-related factors.[6] This correlation illustrates why competency-based medical education (CBME) is critical; we cannot rely on case volume alone. CBME encourages learner-centered graduated levels of trainee responsibility based on outcomes and abilities, which is not synonymous with time in training or volume of prior experience.[7-9] While case volume certainly contributes to a body of surgical experience, that alone does not equate to self-efficacy.

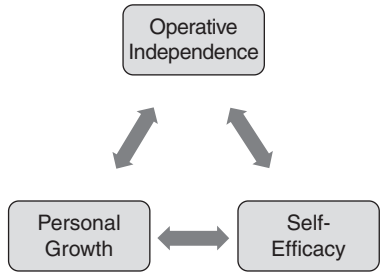

FIGURE
Operative Self-Efficacy Cycle

The comparison of trainee and attending perceptions of operative independence granted is consistent with existing literature suggesting an "autonomy gap" in which residents tend to underrate themselves and perceive needing more supervision compared to attending physician reports.[10] While there is still work to be done to better align trainee/attending perceptions, the findings from the current study suggest that attending-reported supervision level and resident-reported self-efficacy remain strongly correlated. Thus, the lack of operative independence identified in this study cannot simply be explained as a trainee perception issue.

There are several limitations associated with this study. The tiered classification for self-efficacy was used to group cases for statistical analysis purposes instead of considering each operation independently. However, ANOVA did show significant differences between each of the tiers, supporting the notion that

these do represent distinct groups. Additionally, the ACGME case log granularity prevents perfect alignment with the operations chosen for the initial self-efficacy post-ABSITE survey. Thus, the case volume sample size was limited for each of the self-efficacy tiers because several operations had to be excluded from the analysis. There are also varying degrees of difficulty within elements of a single operation; some residents may be facile at a portion of the operation but unable to do the entire case with operative independence, which is not captured by this survey. Similarly, the post-ABSITE self-efficacy survey did not specify the level of case difficulty when asking about operative self-efficacy, which could have led to differences in how the question was perceived by respondents. The use of SIMPL OR also presents a limitation to the generalizability of the study, as not all general surgery programs use the platform.

Future areas of study could seek to design interventions that target opportunities for operative independence and case volume to help facilitate the positive self-efficacy cycle previously described. For cases in the lower self-efficacy tiers, which tend to confer less operative independence and volume, targeted efforts in simulation or video-based education could be considered. Other elements of social persuasion, such as operative feedback and coaching, could also be targeted.

Conclusions

This study shows a strong correlation between operative independence and self-efficacy. Operative independence confers a high degree of resident entrustment through social persuasion, a key tenant of self-efficacy theory. However, case volume or quantity, which address the role of mastery experience in contributing to perceptions of self-efficacy, did not have a significant effect.

References

1. Bandura A, Freeman WH, Lightsey R. Self-efficacy: the exercise of control. *J Cogn Psychother*. 1999;13(2):158-166. doi:10.1891/0889-8391.13.2.158
2. Anderson TN, Payne DH, Dent DL, Kearse LE, Schmiederer IS, Korndorffer JR. Defining the deficit in US surgical training: the trainee's perspective. *J Am Coll Surg*. 2021;232(4):623-627. doi:10.1016/j.jamcollsurg.2020.11.029
3. Kearse LE, Schmiederer IS, Dent DL, et al. Individual and institutional factors associated with PGY5 general surgery resident self-efficacy: a national survey. *J Am Coll Surg*. 2022;234(4):514-520. doi:10.1097/XCS.0000000000000090
4. 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. doi:10.1016/j.jsurg.2014.06.018
5. Accreditation Council for Graduate Medical Education. Surgery Case Logs National Data Report. Accessed April 16, 2024. https://apps.acgme.org/ads/Public/Reports/CaselogNationalReportDownload?specialtyId=99&academicYearId=26
6. Hauer KE, ten Cate O, Boscardin C, Irby DM, Iobst W, O'Sullivan PS. Understanding trust as an essential element of trainee supervision and learning in the workplace. *Adv Health Sci Educ Theory Pract*. 2014;19(3):435-456. doi:10.1007/s10459-013-9474-4
7. Frank JR, Snell LS, ten Cate O, et al. Competency-based medical education: theory to practice. *Med Teach*. 2010;32(8):638-645. doi:10.3109/0142159X.2010.501190
8. Edgar L, Roberts S, Yaghmour NA, et al. Competency crosswalk: a multispecialty review of the Accreditation Council for Graduate Medical Education Milestones across four competency domains. *Acad Med*. 2018;93(7):1035-1041. doi:10.1097/ACM.0000000000002059
9. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The role of assessment in competency-based medical education. *Med Teach*. 2010;32(8):676-682. doi:10.3109/0142159X.2010.500704
10. Kim GJ, Clark MJ, Meyerson SL, et al. Mind the gap: the autonomy perception gap in the operating room by surgical residents and faculty. *J Surg Educ*. 2020;77(6):1522-1527. doi:10.1016/j.jsurg.2020.05.023

Rachel M. Jensen, MD, MHPE, is a PGY-4 Resident, Prior Surgical Education Fellow, Department of Surgery, University of California San Diego, San Diego, California, USA; **Ananya Anand, MD,** is a PGY-3 Resident, Surgical Education Fellow, Department of Surgery, Stanford University, Stanford, California, USA; **LaDonna E. Kearse, MD,** is a PGY-4 Resident, Prior Surgical Education Fellow, Department of Surgery, Mayo Clinic-Rochester, Rochester, Minnesota, USA; and **James R. Korndorffer Jr, MD, MHPE,** is Vice Chair of Education, Department of Surgery, Stanford University, Stanford, California, USA.

Funding: The authors report no external funding source for this study.

Conflict of interest: The authors declare they have no competing interests.

This was presented as an oral presentation at the Accreditation Council for Graduate Medical Education Annual Educational Conference, February 23-25, 2023, Nashville, Tennessee, USA.

The authors would like to thank the Procedural Learning and Safety Collaboration (Boston, Massachusetts, USA) for providing SIMPL OR data.

Corresponding author: Rachel M. Jensen, MD, MHPE, Stanford University, Stanford, California, USA, rachel.m.jensen@gmail.com, X @racheljensenMD

Received July 1, 2023; revisions received December 5, 2023, and April 1, 2024; accepted April 5, 2024.