Quantitative Analysis of Surgical Residency Reform: Using Case-Logs to Evaluate Resident Experience

Sarah B. Cairo, MD, MPH,[*,†] Wendy Craig, PhD,[‡] Caitlin Gutheil, MS,[‡] Paul K.J. Han, MD, MA, MPH,[‡,§] Kristiina Hyrkas, PhD,[¶] Lynda Macken, PhD,[¶] and James F. Whiting, MD[*,‖]

[*]Maine Medical Center Department of Surgery, Portland, Maine; [†]Women and Children's Hospital of Buffalo, Buffalo, New York; [‡]Center for Outcomes Research and Evaluation (CORE) and Maine Medical Center Research Institute, Portland, Maine; [§]Palliative Medicine, Hospice of Southern Maine, Scarborough, Maine; [‖]Clinical Associate Professor of Surgery, Tufts University School of Medicine at Maine Medical Center, Portland, Maine; and [¶]Center for Nursing Research and Quality Outcomes, Maine Medical Center, Portland, Maine

BACKGROUND: Curricular changes at a mid-sized surgical training program were developed to rebalance clinical rotations, optimize education over service, decrease the size of service teams, and integrate apprenticeship-type experiences. This study quantifies the operative experience before and after implementation as part of a mixed-methods program evaluation.

STUDY DESIGN: Retrospective review of case-log data and data from the Accreditation Council for Graduate Medical Education (ACGME) and the American College of Surgeons National Surgical Quality Improvement Program: quality in-training initiative to evaluate case volume pre- and postintervention.

RESULTS: 11,365 cases, excluding "first-assistant" and "endoscopic" cases, were logged for an average of 291 and 263 cases/resident pre- and postintervention, respectively. Average case volume increased significantly for postgraduate year (PGY) 3 residents and decreased significantly for PGY 4 residents between the two time periods. Variability was observed among residents at the same PGY level both pre- and postintervention, with coefficients of variation of 6.0% to 34.1% in 2014 to 2015 and 11.2% to 66.8% in 2015 to 2016. Inter-resident variability persisted when comparing a specific procedure between ACGME case-log and quality in-training initiative data sets.

CONCLUSION: The data suggest that inter-resident variability in case load is not an artifact of case logging behavior alone, but may reflect personal preferences and choices in case selection that are not impacted by curriculum change. Logging behavior and accuracy of case-logs may contribute to variability. The shift in case load from PGY 4 to PGY 3 after curriculum implementation requires validation by ongoing analysis of ACGME case-log data. (J Surg Ed 76:25–35. © 2018 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: Surgical education, Curriculum reform, Case-log, Residency reform, Resident education

COMPETENCIES: Patient Care

BACKGROUND

Surgical education has faced a number of major challenges over the past half century ranging from increased subspecialization, changes in patient expectations, rapid innovations in technology and surgical technique, and duty hours restrictions.[1–3] These changes have prompted a number of proposals and initiatives aimed at redesigning surgical training. Initiatives such as the American College of Surgeon's "Fix the Five," and the American Board of Surgery's "Flexibility in Surgical Training," and "Early Specialization" pathways have emphasized changes in the external structure of training with an as yet unknown effect on the resident experience.[4–6] One barrier to curricular redesign and changing

Correspondence: Inquiries to Sarah B Cairo, MD, MPH, Women and Children's Hospital of Buffalo, 208 Summer St Apt #4, Buffalo, New York 142222; e-mail: scairo2@gmail.com

Presentation: Podium presentation at New England Surgical Society 97th Annual meeting, Bretton Woods, NH, 9/2017.

Funding: This research was funded by Educational Innovations Grants Program through Maine Medical Center and Tufts University School of Health.

  1931-7204/$30.00 https://doi.org/10.1016/j.jsurg.2018.05.013

the model for surgical training lies in the difficulty of defining and measuring success.

We recently redesigned the curriculum in the general surgery residency at our institution with the goals of: (1) optimizing education over service; (2) decreasing the size of service teams; (3) integrating apprenticeship-type experiences into residency; and (4) rebalancing the clinical rotations to reflect the educational goals of the American Board of Surgery In-Service Examination. In evaluating such a curriculum change, the appropriate metrics for success need to address the goals of a curriculum redesign and its effects on the residents' operative experience. We did not set out to "redistribute" cases between postgraduate year (PGY) years or preferentially alter or enhance the operative experience of one group over another. We did anticipate, however, that the change would open up avenues for senior residents, other than those in the PGY 5 position, to develop autonomy in the operating room.

As part of a larger mixed-methods program evaluation, we conducted a study to quantify the operative experience of categorical surgical residents before and after implementation of curricular reform using case-log data to evaluate the impact of the curriculum change on operative volume. Case log data was evaluated as a proxy for surgical proficiency and as a commonly employed metric of success for a surgical program despite the increasing use of competency-based evaluations. Since case-log data rely on the accuracy of entry by individual residents, we also compared American College of Surgeons National Surgical Quality Improvement Program: quality in-training initiative (NSQIP-QITI) data for selected procedures to assess the validity of using case data for program evaluation in the 12 month period prior to and after the curriculum change. We hypothesized that there would not be a statistically or practically relevant change in operative experience (as defined by case-logs) secondary to the curriculum change.

METHODS

The study protocol was reviewed by the Institutional Review Board of Maine Medical Center and a waiver of consent was issued for nonresearch (IRB # 4793NR).

Curriculum Reform

The surgical residency at Maine Medical Center has long consisted of service groups organized by attending surgeons and major categories of surgical care. In 2015 a team of surgical trainees and attending surgeons reviewed the American Board of Surgery map for the American Board of Surgery In-service Trainee Examination (ABSITE) and the Surgical Council on Resident Education curriculum. The various modules within the American Board of Surgery (ABS) map specifically and proportionally reflect the content of the ABSITE. The modules were divided into categories relating to operative experiences which were subsequently organized by proposed surgical team to achieve a sense of balance for the topics residents are responsible for. In contrast to the larger service groups, the new surgical teams were arranged by educational experience as demonstrated in the example shown in Figure 1. Other services which were restructured in this effort include the general surgery service (divided into elective general surgery and acute care surgery), the former "yellow" surgery service (divided into bariatric/minimally invasive surgery and colorectal surgery), and vascular surgery (divided into two services).

Case Selection

This study was conducted at a university-affiliated community program with 4 designated residents per year. Residents operate within a single facility which is the only ACS-verified Level 1 Trauma Center and Pediatric Emergency Department in the state. Resident and facility information were accessed from the Accreditation Council for Graduate Medical Education (ACGME) case-log and the American College of Surgeons (ACS) NSQIP-QITI as part of a multifaceted program wide evaluation and quality improvement initiative. Data were collected for categorical residents from 2014 to 2016 with exclusion of cases with resident experience classified as "first-assistant" or "endoscopic."

ACGME Case-Logs

The first portion of this study involved retrospective review of the ACGME case-log data. These data are entered on a continual basis by general surgery residents enrolled in North American training programs. Data entry includes common procedural terminology codes, date of procedure, resident post graduate year, resident role in case, attending surgeon, surgical team, and pediatric versus adult classification. Cases were evaluated by surgical specialty, resident postgraduate year (PGY), and resident role, as defined by the ACGME as "teaching assistant," where a more senior resident works with a junior resident who is credited as the surgeon; "surgeon junior," including all cases credited as surgeon prior to chief year; and "surgeon chief year," with only cases credited as surgeon during 12 months of chief year.[7] "First assistant" cases, defined as any instance in which a resident assists at an operation with another surgeon (attending or more senior resident) responsible for the operation, were excluded. Case volumes were compared at the individual resident level to assess for variability within PGY and also using PGY as a unit of measure where the average cases for the cohort of interest were included.

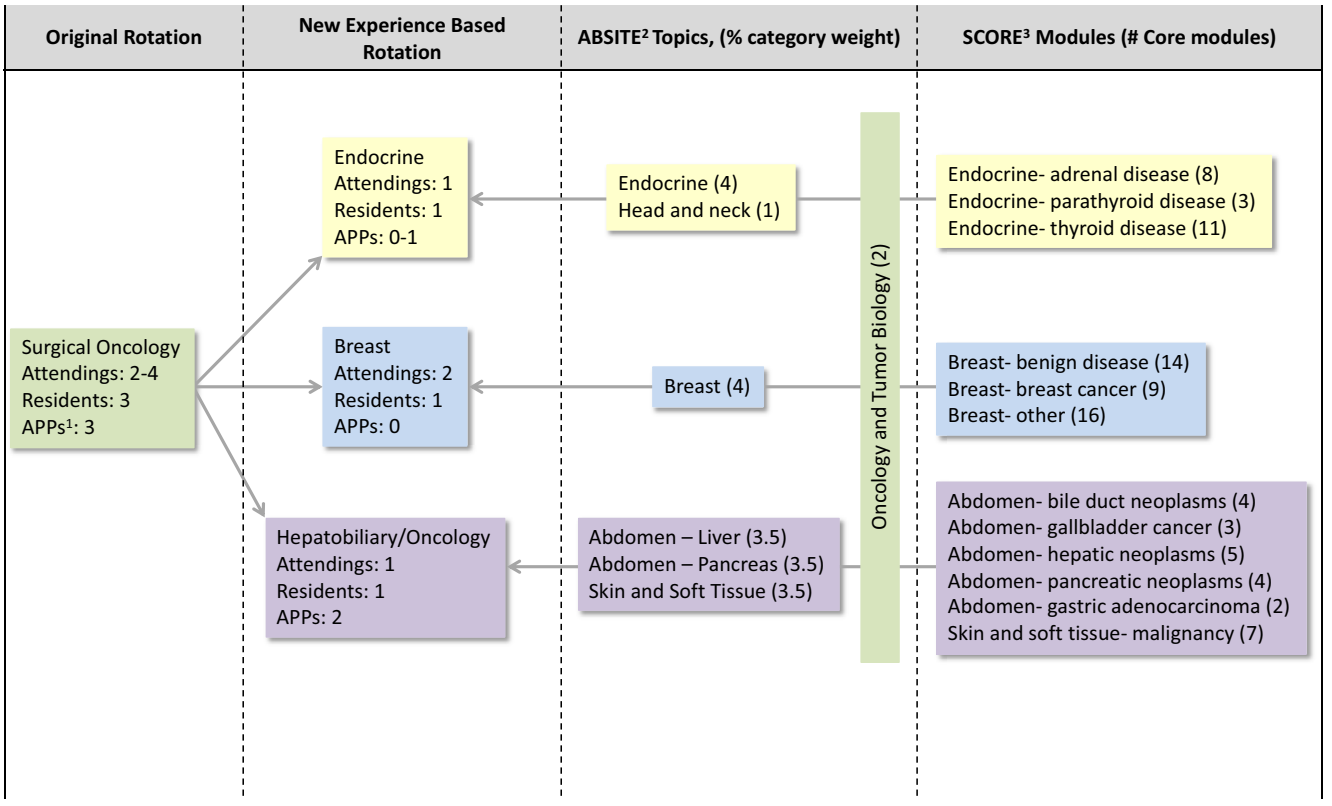

FIGURE 1. Example of changes to surgical services with curriculum reform.
[1]Advanced practice providers: nurse practitioners, physician assistants.
[2]American board of surgery in-service examination.
[3]Surgical council on resident education.

QITI Case-Logs

The second portion of this study utilized the American College of Surgeons NSQIP-QITI. QITI was launched in 2011 with three main objectives: (1) allow analysis of ACS NSQIP outcomes data to provide a benchmark for surgical education; (2) develop a national quality improvement curriculum based on real issues in surgical care; and (3) foster a culture of surgeon leaders with a quality conscience. The QITI utilizes the ACS NSQIP to collect resident-specific data and generate outcome reports.[8] With data abstraction from ACS NSQIP data entered by chart abstractors, QITI provides granular data used by program directors to satisfy ACGME requirements for quality improvement.[9] Resident report case details include principal procedure based on common procedural terminology code, date of operation, NSQIP case number, attending surgeon, comorbidities (total number), and list of complications. While the majority of cases are identified for inclusion in NSQIP and QITI based on a systematic sampling system, select procedures at our institution have been identified for 100% sampling (colectomy and appendectomy). We used these cases as a comparison to ACGME case-log data.

Statistical Analysis

Data were summarized as mean case number per resident, stratified by PGY and by study period (2014 to 2015 and 2015 to 2016), both for overall surgical cases (excluding endoscopy and first assists) and for selected procedures with data available from both the ACGME case-logs and from QITI. Differences in residents' case numbers (from ACGME case-log data) before and after curriculum reform, stratified by PGY, were evaluated by *t* tests. Variability in case numbers within a PGY was assessed as coefficient of variation (%CV). Paired Wilcoxon signed rank test was utilized to assess variability between the two case-log databases for individual residents. All data analysis was performed with Microsoft Excel (2010). Significance was accepted at $p < 0.05$.

RESULTS

ACGME case-logs were accessed for all categorical residents at a single institution preintervention (2014 to 2015) and postintervention (2015 to 2016). A total of 11,365 cases, excluding "first-assistant" and

TABLE 1. ACGME Log Case Distribution for Total Surgeries,* Before and After Curriculum Change, by Postgraduate Year

PGY	2014 to 2015 (Precurriculum Change)						2015 to 2016 (Postcurriculum Change)						
	Trainees (n)	Total Cases (n)	Cases/Resident Mean	Min	Max	%CV	Trainees (n)	Total Cases (n)	Cases/Resident Mean	Min	Max	%CV	p Value†
1	4	375	94	55	129	34.1	4	396	99	35	190	66.8	p = 0.89
2	4	978	244	218	318	20.1	4	746	186	146	237	24.7	p = 0.14
3	4	1047	262	250	284	6.0	4	1413	353	301	382	11.2	**p = 0.005**
4	4	1469	367	340	424	11.1	4	1131	283	253	341	17.2	**p = 0.04**
5	5	2241	448	303	551	24.1	4	1569	392	361	480	15.1	p = 0.39
Total	21	6110	291	55	551	-	20	5255	263	35	480	-	-

Bold: statistical significance, $p < 0.05$.
Abbreviations: %CV = percent coefficient of variation; min = minimum; max = maximum; PGY = postgraduate year.
*Total cases excluding first assists and endoscopy.
†*t* test was used to compare number of cases/resident within each PGY, pre- and postcurriculum change.

"endoscopic" cases were logged for an average of 291 and 263 cases/resident pre- and postintervention, respectively (Table 1). The total number of cases logged decreased from 6110 prior to curriculum change to 5255 postintervention; however, the resident complement also decreased by 1 during that time period. Variability in case numbers within each PGY, expressed as %CV, or the ratio of the standard deviation to the mean, was 6.0% to 34.1% preintervention and 11.2% to 66.8% postintervention. The average number of cases logged per resident increased significantly for residents in the third postgraduate year (262 cases per resident to 353 cases per resident, respectively, p = 0.005). Conversely, the average number of cases decreased significantly for residents in the fourth postgraduate year (367 cases per resident to 283 cases per resident, respectively, p = 0.037). There was no significant difference in average cases per resident between pre and postcurriculum reform for other postgraduate years.

Total cases were then stratified by PGY and by resident role, as determined by ACGME case-log data entered by the resident. With regard to resident role, the total number of cases logged as surgeon junior compared to surgeon senior and teaching assistant were compared for each resident level between the 2 years evaluated (Table 2). Variability was evaluated for statistical significance for exploratory purposes despite small sample size and there were no statistically significant differences between average case distribution across resident classes for surgeon chief (p = 0.73), surgeon junior (p = 0.54), or teaching assistant (p = 0.97). There were nonsignificant trends toward more cases logged as surgeon chief and teaching assistant for PGY 4 and PGY 5 residents, and toward more surgeon junior cases for junior and mid-level residents (PGY 1-2 and PGY 3-4, respectively).

Specific procedures were selected for more detailed analysis based on anticipated change in case distribution with the curriculum reform and major cases required for graduation including colectomy and appendectomy (Table 3). Total pre- and postintervention data were compared as well as average number of cases per resident at each PGY level. There was a significant difference in total number of colectomies by PGY level (p = 0.03), but no other differences in resident role by PGY level before and after curriculum change for the procedures evaluated. Surgery type was also evaluated across PGY and pre- and postintervention with most pronounced change for abdomen total, surgical critical, endocrine, and plastic surgery (Fig. 2).

Further analysis was performed to compare ACGME case-logs and QITI data for procedures captured fully by both data sets. Specifically, colectomy and appendectomy were assessed for pre- and postintervention changes. For colectomies, there were no significant differences in mean number of cases by PGY before and after curriculum change (Table 4). Variability in case numbers within each postgraduate year, expressed as %CV, was greatest for the PGY 1 and 2 level residents for both ACGME and QITI data. For ACGME case-log-data, %CV was greater postcurriculum change for all PGY levels except PGY 1. For QITI data, %CV was greater postcurriculum change for PGY 2 level residents but greater precurriculum change for all others. Similarly, the greatest %CV was observed for PGY 1 residents in both datasets for appendectomy (Table 5). For QITI data, %CV was greater postcurriculum change at all PGY levels. For ACGME data, however, %CV was greater postcurriculum change for all PGY levels except PGY 3 where a greater %CV was observed precurriculum change.

In order to assess reliability between the two datasets, Wilcoxon signed rank test was utilized to compare the data at the individual resident level. Colectomy was used as an example as the procedure is captured fully by both the ACGME case-logs and QITI. For the first year, 2014

TABLE 2. Distribution of ACGME Case Log Data, Stratified by Resident Roles, by PGY and by Pre- and Postintervention Time Period*

Resident Role	Pre/Post†	PGY 1	PGY 2	PGY 3	PGY 4	PGY 5	Total
Surgeon Chief	Pre	2 (0.5) [0.5]	3 (0.8) [0.3]	0 (0.0) [0]	3 (0.8) [0.2]	2035 (407.0) [90.8]	2043 (97.3) [33.4]
	Post	5 (1.3) [1.3]	20 (5.0) [2.7]	1 (0.3) [0.1]	360 (90.0) [31.8]	1418 (354.5) [90.4]	1804 (90.2) [34.3]
Surgeon Junior	Pre	371 (92.8) [98.9]	949 (237.3) [97.0]	1040 (260.0) [99.3]	1421 (355.3) [96.7]	85 (17.0) [3.8]	3866 (184.1) [63.3]
	Post	390 (97.5) [98.5]	721 (180.3) [96.6]	1398 (349.5) [98.9]	744 (186) [65.8]	23 (5.8) [1.5]	3276 (163.8) [62.3]
Teaching Assistant	Pre	2 (0.5) [0.5]	26 (6.5) [2.7]	7 (1.8) [0.7]	45 (11.3) [3.1]	121 (24.2) [5.4]	201 (9.6) [3.3]
	Post	1 (0.3) [0.3]	5 (1.3) [0.7]	14 (3.5) [1.0]	27 (6.8) [2.4]	128 (32.0) [8.2]	175 (8.8) [3.3]
Total	Pre	375 (93.8) [6.1]	978 (244.5) [16.0]	1047 (261.8) [17.1]	1469 (367.3) [24.0]	2241 (448.2) [36.7]	6110 (291.0)
	Post	396 (99.0) [7.5]	746 (186.5) [14.2]	1413 (353.3) [26.9]	1131 (282.8) [21.5]	1569 (392.3) [29.9]	5255 (262.8)

Abbreviation: PGY = postgraduate year.
* Results are displayed for each year, with case numbers presented overall and after stratification by resident role and by PGY. The data are shown for each time period as: total number of cases per PGY/role category [average number of cases per resident, based on number of trainees in the specified PGY/role category] [percentage of total cases performed within each resident role, overall and by PGY].
† Pre: 2014 to 2015, Post: 2015 to 2016.

to 2015, the mean difference between QITI and ACGME case-logs was −2.2 cases; in 2015 to 2016, the difference was −2.4 cases with QITI lower than ACGME case-logs for both years. There was no significant difference between the two measures of assessment for total colectomies performed by residents prior to curriculum reform ($p = 0.19$) or after reform ($p = 0.15$).

DISCUSSION

In the wake of an evolving healthcare system with increased demands on surgical trainees and constraints on educational freedom, the creation of an efficient training program is critical. Alongside curriculum reform with a framework in adult education theory, mechanisms to evaluate resident performance should reliably assess trainees for competency at variable levels.[10] Despite a strong trend toward competency-based evaluations, case volume remains a metric against which surgical proficiency and success at the individual and programmatic level are gauged. Following surgical residency reform at a single institution and implementation of competency based evaluations nationwide, this study set out to evaluate the impact of curricular changes on resident operative experience while evaluating the use of ACGME case-logs and QITI case records for individual and institutional assessment.

The American Board of Surgery (ABS) was established in 1937 to develop a certification process and national certifying body for surgeons practicing in the United States. Operative experience is part of the assessment of trainee eligibility with a minimum of 850 operative procedures being performed in the 5 years of residency with 200 procedures in the chief resident year and 250 cases by the end of the second postgraduate year.[11] These cases are recorded by trainees and submitted to the ACGME and the ABS for review at the individual and program level.[12]

Increased emphasis on patient safety, accountability, and surgical performance has resulted in work hour reform and consequently, variable reports of the effect of these changes on the trainee experience.[13–15] In the wake of these changes or restrictions with increasing minimum operative experience requirements of the ABS, the process of attaining surgical expertise has been called into question. Plainly stated, experience, though essential to becoming an expert, does not necessarily equate to expertise and therefore the number of surgical procedures or periods of deliberate practice, are not as clearly defined as once proposed.[16–18] Further supporting the flawed utility of procedure volume alone, surveys of program directors, fellowship directors, and graduating surgical residents alike have failed to identify a

TABLE 3. Resident Roles Stratified by PGY and Pre- and Postintervention Data by Major Case Type, ACGME Case Log

Resident Role	Pre/Post	PGY 1*	PGY 2	PGY 3	PGY 4	PGY 5	Total	p Value†
(a) Colectomy								
Surgeon Chief	2014 to 2015	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.3)	185 (37.0)	186 (8.9)	p = 0.17
	2015 to 2016	0 (0.0)	0 (0.0)	1 (0.3)	30 (7.5)	177 (44.3)	208 (10.4)	
Surgeon Junior	2014 to 2015	1 (0.3)	14 (3.5)	25 (6.3)	112 (28.0)	5 (1.0)	157 (7.5)	p = 0.82
	2015 to 2016	9 (2.3)	16 (4.0)	36 (9.0)	100 (25.0)	0 (0.0)	161 (8.1)	
Teaching Assistant	2014 to 2015	0 (0.0)	0 (0.0)	1 (0.3)	2 (0.5)	7 (1.4)	10 (0.5)	p = 0.08
	2015 to 2016	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.3)	5 (1.3)	6 (0.3)	
Total	2014 to 2015	1 (0.3)	14 (3.5)	26 (6.5)	115 (28.8)	197 (39.4)	353 (16.8)	p = 0.03
	2015 to 2016	9 (2.3)	16 (4.0)	37 (9.3)	131 (32.8)	182 (45.5)	375 (18.8)	
(b) Appendectomy								
Surgeon Chief	2014 to 2015	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	37 (7.4)	37 (1.8)	p = 0.71
	2015 to 2016	0 (0.0)	0 (0.0)	0 (0.0)	9 (2.3)	25 (6.3)	34 (1.7)	
Surgeon Junior	2014 to 2015	37 (9.3)	47 (11.8)	77 (19.3)	80 (20.0)	3 (0.6)	244 (11.6)	p = 0.20
	2015-2016	35 (8.8)	43 (10.8)	69 (17.3)	40 (10.0)	0 (0.0)	187 (9.4)	
Teaching Assistant	2014 to 2015	0 (0.0)	0 (0.0)	3 (0.8)	11 (2.8)	3 (0.6)	29 (1.4)	p = 0.72
	2015 to 2016	0 (0.0)	0 (0.0)	0 (0.0)	13 (3.3)	15 (3.8)	28 (1.4)	
Total	2014 to 2015	37 (9.3)	47 (11.8)	80 (20.0)	91 (22.8)	55 (11.0)	310 (14.8)	p = 0.12
	2015 to 2016	35 (8.8)	43 (10.8)	69 (17.3)	62 (15.5)	40 (10.0)	249 (12.5)	

Abbreviation: PGY = postgraduate year.

*Total per PGY (average number of cases per resident based on number of trainees per PGY per year).

†*t* test for statistical significance.

correlation between robust case volumes and competence.[19,20] Regardless of the validity of using procedure numbers as a proxy for surgical experience or expertise, the system remains heavily dependent on these case-logs as a metric for success.[21] Resident role, an adjunct measure to evaluate resident autonomy and advancing level of expertise, did not appear to change significantly with the curriculum change in our study. More detailed assessment, including evaluation of specific case types and use of live feedback tools describing resident involvement in various stages of a surgical procedure, are needed to quantify the effect of the curriculum change on resident autonomy.

In this study, we identified more variability in case log experience between residents of the same year than was introduced by structural curriculum change. Variability was assessed using range of cases performed per resident within a specific category and overall and using %CV expressed as the ratio of the standard deviation to the mean. In other words, the %CV was used to measure the relative variability of the data sets on a ratio scale. While the sample size was too small to determine true variance or statistical significance of the variance between members of the same PGY level in each dataset before and after curriculum change, descriptive data were obtained using these simple calculations. In general higher %CVs were observed in the postcurriculum change group for total cases, colectomy, and appendectomy. Higher %CVs were also noted in residents at a lower PGY level; this was more pronounced in the ACGME case-logs compared to QITI data.

While individual resident case-log behavior is thought to account for some irregularities between residents of the same PGY level within this study, other studies have identified increased variability in resident case volume following the implementation of the 80-hour workweek.[22] In one such study, Quillin et al. reported increased range or variability between the resident finishing with the most and the fewest cases.[23] It is noted that the change was primarily observed as a significant decrease in the number of cases performed by the resident with the fewest cases. Differences in learner type, likelihood of attaining a postgraduate fellowship, and resident attrition have also been linked to variability in operative volume.[24]

In contrast to the ACGME case-logs where significant variability was observed, the variability at the PGY 1 level dissipated when analyzing the data from QITI. In contrast to the ACGME case-logs which, in most institutions, depend on the individual resident to enter data, QITI generates custom reports for ACGME-accredited surgical training programs using data from the ACS-NSQIP as an educational tool to teach residents about general outcome measures and their utility in quality improvement.[25,26] The data, therefore, are collected by a trained surgical clinical reviewer and should not vary between surgical residents. Some of the observed inter-resident variability may have been attributable to variable understanding of the resident roles specified by the ACGME.[27,28] In a survey of surgical residents in a large academic medical center, trainees reported perceived lack of accuracy of case-logs with nearly 48.1% of

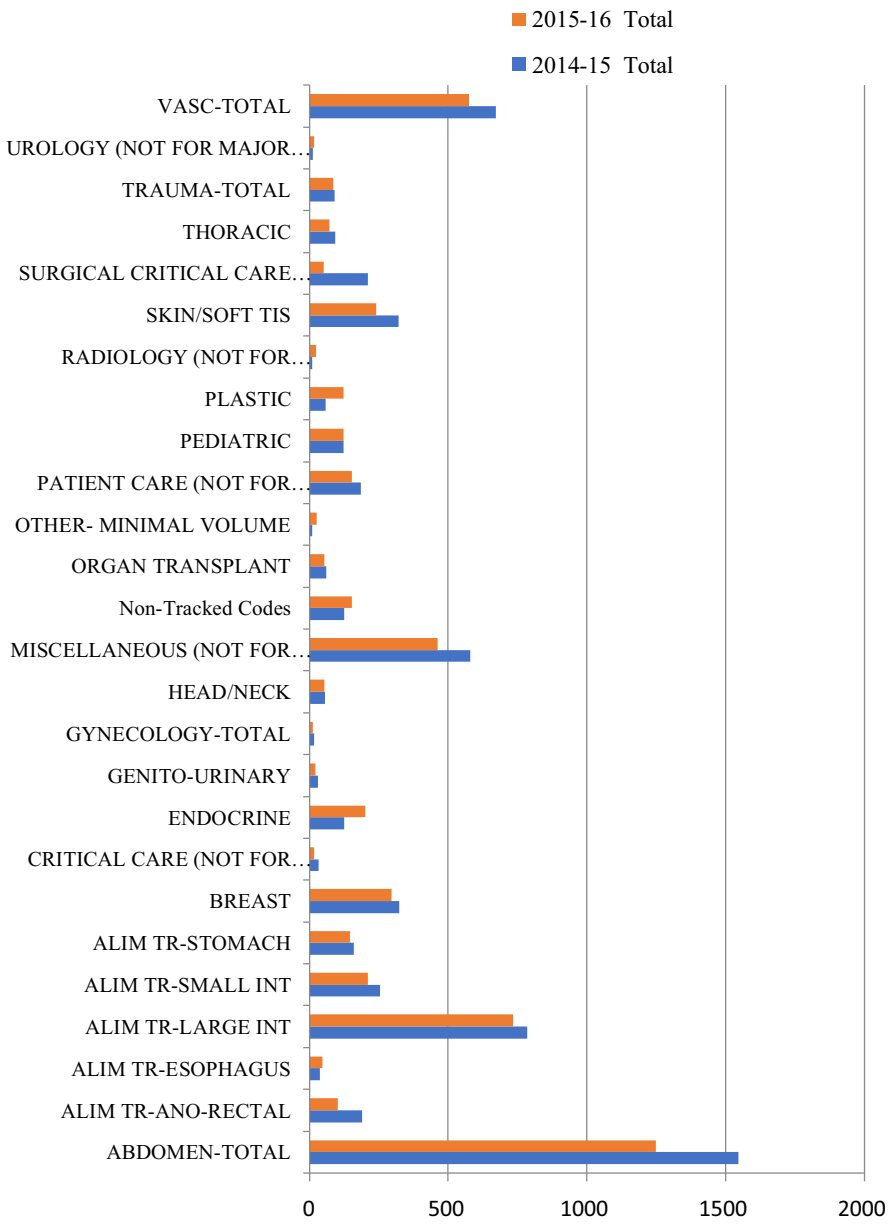

FIGURE 2. Total cases by surgical category pre- and postintervention.

trainees using incorrect metrics to log their role as surgeon or assistant.[29] Our case logs demonstrated similar errors in accuracy of assigning resident role such as cases logged as "surgeon chief" by residents other than PGY 5s (this would require prior ACGME approval which we do not have). These inconsistencies are an example of the variability in logging behavior. Subsequent to this study, residents have been re-educated on the existing ABS definitions of surgeon roles in operating case-logs. Frequency of errors in case-logs and need for clarification to inherent vagueness in role definitions (i.e. difference between first assist and surgeon junior, perioperative planning requirements for logging a case a surgeon, etc) are being addressed through semiannual review and ongoing research..While errors in case-log behavior likely account for some of the variability seen in the ACGME case-logs at the PGY 1 level, the variability in later years is more likely associated with resident preference and career choice. In our program, cases are generally assigned by the most senior resident involved, and although attendings can request a certain level of resident to be present on a case, that prerogative is rarely invoked. The senior resident also determines what cases are "uncovered" or covered by an advanced practice professional when there are more cases than residents to be done. Though speculative in nature, the concept

TABLE 4. Case Distribution for Colectomy from ACGME and QITI Datasets, Stratified by PGY and Pre- and Postintervention

PGY	2014 to 2015 (Precurriculum Change)						2015 to 2016 (Postcurriculum Change)						
	Trainees	Cases	Cases/Resident				Trainees	Cases	Cases/Resident				
	(n)	(n)	Mean	Min	Max	%CV	(n)	(n)	Mean	Min	Max	%CV	p Value*
(a) Colectomy (ACGME Case Log Data)													
1	4	1†	0.3	0	1	200.0	4	9‡	2.3	0	7	146.8	p = 0.28
2	4	14	3.5	1	7	71.9	4	16‡	4	0	12	141.4	p = 0.88
3	4	26	6.5	4	11	47.8	4	37	9.3	5	16	31.2	p = 0.39
4	4	115	28.8	20	37	24.3	4	131	32.8	20	41	26.3	p = 0.54
5	5	197	39.4	29	55	26.3	4	182	45.5	30	67	39.8	p = 0.54
(b) Colectomy (QITI data)													
1	4	5§	1.3	0	3	100.7	4	8§	2.0	0	3	70.7	p = 0.46
2	4	5§	1.3	0	3	100.7	4	30‡	7.5	0	20	127.7	p = 0.24
3	4	30	7.5	5	10	31.7	4	16	4.0	1	8	73.6	p = 0.11
4	4	87	21.8	17	30	27.2	4	116	29.0	21	36	23.4	p = 0.16
5	5	184	36.8	22	58	37.0	4	172	43.0	30	51	22.7	p = 0.94

* Student's *t* test to compare mean pre- and postcurriculum change.
† 3 of 4 residents performed 0 cases.
‡ 2 of 4 residents performed 0 cases.
§ 1 of 4 residents performed 0 cases.

of individualized resident experience based on personal preference warrants more research, especially in the era of global surgical residency restructuring. The variability between members of the same PGY level was observed before and after curriculum reform suggesting that curriculum structure may be less influential on overall operative experience than personal preference.

As observed by Quillin et al. with regard to overall case volumes, learner type and likelihood of attaining a postgraduate fellowship appeared to influence resident experience before and after curriculum reform.[24,29] Residents entering into general surgery practice, for example, were noted to have significantly higher numbers of alimentary tract cases compared to their colleagues continuing on to surgical specialty training. For example, a senior resident in the cohort studied, planning on a career in rural general surgery, performed five times as many inguinal hernias as any of his classmates as a senior resident. Having not formally compared the operative experience of individual residents by career preferences prior to this study, we could not have anticipated the variability observed between residents and had not sought to address standardization of the resident experience other than to adequately expose residents to the essential surgical experiences. Based on career aspirations and operative competency, however, it is unclear

TABLE 5. Case Distribution for Appendectomy from ACGME and QITI Datasets, Stratified by PGY and Pre- and Postintervention

PGY	2014 to 2015 (Precurriculum Change)						2015 to 2016 (Postcurriculum Change)						
	Trainees	Cases	Cases/Resident				Trainees	Cases	Cases/Resident				
	(n)	(n)	Mean	Min	Max	%CV	(n)	(n)	Mean	Min	Max	%CV	p Value*
(a) Appendectomy (ACGME Case Log Data)													
1	4	37	9.3	3	16	58.1	4	35	8.8	2	22	102.8	p = 0.93
2	4	47	11.8	10	16	24.4	4	43	10.8	5	15	43.2	p = 0.73
3	4	80	20.0	13	32	41.4	4	69	17.3	11	23	28.5	p = 0.59
4	4	91	22.8	17	33	31.4	4	62	15.5	6	24	47.6	p = 0.21
5	5	55	11.0	8	15	26.5	4	40	10.0	3	15	50.3	p = 0.72
(b) Appendectomy (QITI data)													
1	4	22	5.5	3	10	52.2	4	24	6.0	1	10	81.2	p = 0.93
2	4	28	7	4	12	44.0	4	26	6.5	2	15	77.3	p = 0.87
3	4	37	9.3	5	14	37.7	4	24	6.0	3	11	51.4	p = 0.24
4	4	58	14.5	12	17	13.0	4	42	10.5	7	17	37.2	p = 0.10
5	5	42	8.4	6	13	28.8	4	32	8.0	4	11	38.5	p = 0.83

* Student's *t* test to compare mean pre- and postcurriculum change.

if this variability is favorable or poses a limitation of a curriculum reformed to reflect the ABS training examination.

An overall decrease in case volume was identified in comparing the pre- and postimplementation ACGME case-logs. A decrease in total complement of surgical residents likely contributes to this change. (Roughly 20% of cases performed by our "core" attendings on general surgery services are either uncovered or covered by Advanced Practice Professionals/First assists who work directly on the resident teams, so a decrease in resident component does not automatically result in an increase in average resident cases). To compensate for this change, average number of cases per resident was utilized. This study was designed in part to ensure that the curriculum reform did not result in a significant decline in case volume or operative experience per resident. We believe that the slight decrease in cases, as observed here, may carry statistical significance despite absence of practical significance.

The change in case logs was significant only for average total cases excluding first assist and endoscopy at the PGY 4 level with a statistically significant increase in average total cases at the PGY 3 level. Given the small sample size available and exploratory nature of this study, these changes may not be directly attributed to the curriculum reform but rather resident logging behavior and operative preference. Additionally, the postreform case volumes are sufficiently greater than the minimum requirements suggesting that while they may be statistically significant across 2 years, they are unlikely to be clinically or practically relevant. With regard to case distribution, variability was identified in total abdominal cases, endocrine, miscellaneous, plastic surgery, and surgical critical care. While some of the changes are statistically significant in exploratory analysis, it is the authors' opinion that they correspond with changes in distribution of cases as per restructuring of the curriculum and services in addition to variable resident case-log behavior as described supra. Surgical critical care, for example, was not affected by the change in curriculum despite an apparent drop in case volume. This is attributed to resident case-log behavior. Breast surgery, on the other hand, was found to have significant changes in quality of cases and educational experience as assessed through the mixed methods program evaluation though no significant change was observed in overall case volume.

Limitations

While this study successfully explored the surgical experience of residents before and after curriculum reform, there are several limitations worth noting. This study was conducted at a single institution with a small number of residents and limited pre- and postimplementation data. The small sample size not only limits generalizability of specific results but also increases the risk of type II error such that statistical significance in average number of cases per residents may appear falsely significant.

The concurrent implementation of several changes limits the ability to determine the specific causes of the observed outcomes. In addition to restructuring the surgical teams, faculty and other structural changes likely influenced the resident experience. Further research including the correlation of intraoperative competency tools to assess resident experience over the course of an intensive, immersion type rotation, for example, would greatly aid in evaluating the effect of smaller more focused rotations. Additionally, statistical correlation between the operative and clinical experience in the new model and score on relevant sections of the ABSITE would support the curriculum reform.[30]

The reliance on individual attending presence and schedules for the smaller, apprenticeship like rotations, a component of the curriculum reform, was not specifically addressed in this study. For the endocrine surgery rotation, for example, one surgeon performed the majority of procedures. While beneficial to trainees in providing continuity and an opportunity for observed growth, this model becomes more complicated when the attending is away or has variability in his/her operating schedule. Contingency plans such as planning ahead to coordinate resident vacations or rotate the resident to an alternate but similar rotation or experience have not yet been implemented. As noted, the curriculum reform was not designed to provide preference to one PGY class over another but rather to enhance the experience of all residents. While smaller services and apprenticeship like rotations may be expected to decrease the variability in resident experience by decreasing the number of residents available to cover specific cases, this was not observed and likely reflects the need for more data given the already small sample size. Ongoing data collection is underway to identify whether these scheduling conflicts and smaller services create significant discrepancies in resident experience or if the high case volume when the surgeon is present are sufficient.

CONCLUSIONS

We set out on this curriculum reform with four goals as delineated in our introduction. This case-log analysis reflects the curriculum goals and supports our hypothesis that there would be no significant impact of a comprehensive curricular reform on resident case-log and operative experience. These findings suggest that resident preference may play a significant role in operative experience, regardless of curriculum structure. Additionally, this study questions the utility of case-log data as a

tool for resident evaluation. Further validation of the operative data comparing ACGME case-logs and QITI is needed to determine the optimal use of different logging systems to increase efficiency and accuracy of case-log records. Ongoing evaluation of these tools together with competency-based metrics and standardization, however, is essential to understanding the utility of procedure logs in evaluating surgical training models. Ultimately the success or failure of the curricular reform undertaken in our program will be a cumulative determination, based on additional metrics including resident and attending satisfaction, ABSITE scores, and Board pass rates.

REFERENCES

1. Bell RH, Biester TW, Tabuenca A, et al. Operative experience of residents in US general surgery programs: a gap between expectation and experience. *Ann Surg*. 2009;249:719–724.
2. Lewis FR, Klingensmith ME. Issues in general surgery residency training—2012. *Ann Surg*. 2012;256:553–559.
3. Klingensmith ME. The future of general surgery residency education. *JAMA Surg*. 2016;151:207–208.
4. Klingensmith ME, Awad M, Delman KA, et al. Early results from the flexibility in surgical training research consortium: resident and program director attitudes toward flexible rotations in senior residency. *J Surg Educ*. 2015;72:e151–e157.
5. Klingensmith ME, Potts JR, Merrill WH, et al. Surgical training and the early specialization program: analysis of a national program. *J Am Coll Surg*. 2016;222:410–416.
6. Debas HT, Bass BL, Brennan MF, et al. American surgical association blue ribbon committee report on surgical education: 2004. *Ann Surg*. 2005;241:1–8.
7. Case Log Information. 2010-2017; http://www.acgme.org/Specialties/Case-Log-Information/pfcatid/24/Surgery, 2017.
8. Sakran JV, Hoffman RL, Ko C, Kelz RR. The ACS NSQIp quality in-training initiative: educating residents to ensure the future of optimal surgical care. *Bull Am Coll Surg*. 2013;98:30–35.
9. Sellers MM, Reinke CE, Kreider S, et al. American college of surgeons NSQIP: quality in-training initiative pilot study. *J Am Coll Surg*. 2013;217:827–832.
10. Rashid P. Surgical education and adult learning: integrating theory into practice. *F1000Res*. 2017;6:143.
11. *Booklet of Information, Surgery 2016-2017*. Philadelphia, American Board of Surgery, Office of the Secretary. PA 2016.
12. About ACGME. 2016.
13. Dennis BM, Long EL, Zamperini KM, Nakayama DK. The effect of the 16-hour intern workday restriction on surgical residents' in-hospital activities. *J Surg Educ*. 2013;70:800–805.
14. Kamine TH, Gondek S, Kent TS. Decrease in junior resident case volume after 2011 ACGME work hours. *J Surg Educ*. 2014;71:e59–e63.
15. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. *Ann Surg*. 2014;259:1041–1053.
16. Kirkman MA. Deliberate practice, domain-specific expertise, and implications for surgical education in current climes. *J Surg Educ*. 2013;70:309–317.
17. Bosse HM, Mohr J, Buss B, et al. The benefit of repetitive skills training and frequency of expert feedback in the early acquisition of procedural skills. *BMC Med Educ*. 2015;15:22.
18. Platz F, Kopiez R, Lehmann AC, Wolf A. The influence of deliberate practice on musical achievement: a meta-analysis. *Front Psychol*. 2014;5:646.
19. Fronza JS, Prystowsky JP, DaRosa D, Fryer JP. Surgical residents' perception of competence and relevance of the clinical curriculum to future practice. *J Surg Educ*. 2012;69:792–797.
20. Mattar SG, Alseidi AA, Jones DB, et al. General surgery residency inadequately prepares trainees for fellowship: results of a survey of fellowship program directors. *Ann Surg*. 2013;258:440–449.
21. Safavi A, Lai S, Butterworth S, Hameed M, Schiller D, Skarsgard E. Does operative experience during residency correlate with reported competency of recent general surgery graduates. *Can J Surg*. 2012;55:S171–S177.
22. Balla F, Garwe T, Motghare P, et al. Evaluating coding accuracy in general surgery residents' accreditation council for graduate medical education procedural case logs. *J Surg Educ*. 2016;73:e59–e63.
23. Quillin RC, Cortez AR, Pritts TA, Hanseman DJ, Edwards MJ, Davis BR. Operative variability among

residents has increased since implementation of the 80-hour workweek. *J Am Coll Surg*. 2016;222:1201–1210.

24. Quillin RC, Pritts TA, Hanseman DJ, Edwards MJ, Davis BR. How residents learn predicts success in surgical residency. *J Surg Educ*. 2013;70:725–730.

25. Kelz RR, Sellers MM, Reinke CE, Medbery RL, Morris J, Ko C. Quality in-training initiative—a solution to the need for education in quality improvement: results from a survey of program directors. *J Am Coll Surg*. 2013;217:1126–1132.e1121-1125.

26. Mrdutt MM, Isbell CL, Regner JL, et al. NSQIP-based quality improvement curriculum for surgical residents. *J Am Coll Surg*. 2017;224:868–874.

27. Perone JA, Fankhauser GT, Adhikari D, et al. Who did the case? Perceptions on resident operative participation. *Am J Surg*. 2017;213:821–826.

28. Morgan R, Kauffman DF, Doherty G, Sachs T. Resident and attending perceptions of resident involvement: an analysis of ACGME reporting guidelines. *J Surg Educ*. 2017;74:415–422.

29. Cadish LA, Fung V, Lane FL, Campbell EG. Surgical case logging habits and attitudes: a multispecialty survey of residents. *J Surg Educ*. 2016;73:474–481.

30. 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.

SUPPLEMENTARY INFORMATION

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jsurg.2018.05.013.