Trends in Operative Case Logs of Chief Residents in Surgery by Sex and Race

A 5-year National Study

Polina Zmijewski, MD, MA,✉ Yoon Soo Park, PhD,†‡ Sean Hogan, PhD,‡ Eric Holmboe, MD,‡ Mary Klingensmith, MD,‡§ Alexander Cortez, MD,‖ Brenessa Lindeman, MD,* Herbert Chen, MD,* Brigitte Smith, MD,¶ and Jessica Fazendin, MD**

Objective: To examine case logs reported by general surgery residents and identify potential disparities in operative experience.

Background: A recent study of 21 institutions noted significant differences between the number of cases reported during general surgery residency by trainees who are underrepresented in medicine (URiM) versus trainees who are not URiM (non-URiM). This study also identified differences between female residents and male residents. We partnered with the Accreditation Council for Graduate Medical Education to examine case logs reported from all accredited general surgery programs in the United States. This is the first time these data have been examined nationally.

Methods: We examined total case logs submitted by graduating residents between 2017 and 2022. Group differences in mean reported case logs were examined using paired *t* tests for female versus male and URiM versus non-URiM overall case numbers.

Results: A total of 6458 residents submitted case logs from 319 accredited programs. Eight-hundred fifty-four (13%) were URiM and 5604 (87%) were non-URiM. Over the 5-year study period, URM residents submitted 1096.95 (SD ± 160.57) major cases versus 1115.96 (± 160.53) for non-URiM residents (difference = 19 cases, $P = 0.001$). Case logs were submitted by 3833 (60.1%) male residents and 2625 (39.9%) female residents over the 5-year study period. Male residents reported 1128.56 (SD ± 168.32) cases versus 1091.38 (± 145.98) cases reported by females (difference = 37.18, $P < 0.001$). When looking at surgeon chief and teaching assistant cases, there was no significant difference noted between cases submitted by URiM versus non- URiM residents. However, male residents reported significantly more in both categories than their female peers ($P < 0.001$).

Conclusions: Overall, URiM residents submitted fewer cases in the 5-year study period than their non-URiM peers. The gap in submitted cases between male and female residents was more pronounced, with male residents submitting significantly more cases than their female counterparts. This finding was consistent and statistically significant throughout the entire study period, in most case categories, and without narrowing of difference over time. A difference of 30 to 40 cases can amount to 1 to 3 months of surgical training and is a concerning national trend deserving the attention of every training program and our governing institutions.

Keywords: case log disparities, case logs, disparities, racial disparities, sex disparities, surgical education, surgical training

(*Ann Surg* 2024;280:473–479)

Graduates of general surgery residencies who wish to be licensed through the American Board of Surgery (ABS) must complete a number of requirements, including passing the Fundamentals of Laparoscopic Surgery examination, the Fundamentals of Endoscopic Surgery examination, the written Qualifying Examination, and the oral Certifying Examination. As a surrogate measure of competency, the ABS also requires that each graduating chief resident submit a case log detailing their operative experiences by year in a number of predetermined categories to the American Council for Graduate Medical Education (ACGME).

Number of cases and exposure to cases of higher complexity have long been considered markers of quality of training in general surgery. Many attending surgeons view the chief resident year as one that is particularly formative, as it is a year of increased autonomy both in patient care and in the operating room. Number and type of cases performed may be affected by a number of factors including, but not limited to: institution type (academic vs community), institution setting (urban vs rural), number of residents per class, and geographic region of the United States. Number and complexity of cases can also be determined by resident workload, patient care support, and attending or resident preference.

Limited research exists in the realm of disparities in operative experience, especially with regard to sex and race. Sex disparities in case numbers have been demonstrated in the robotic colon and rectal surgery,[1] as well as in the field of otolaryngology,[2] with females performing fewer cases during training than males. Two smaller studies within the field of ophthalmology looked at operative case logs[3,4] and also consistently demonstrated a smaller number of cases performed by female trainees. A 2022 single-institution study by Eruchalu et al[5] demonstrated that fewer teaching assistant cases are performed by female residents in general surgery. No small single center studies exist in the literature looking at case numbers and disparities by race.

Recent studies led by the U.S. Resident Operative Experience (ROPE) Consortium, a group of 21 general

From the *Department of Surgery, University of Alabama at Birmingham, Birmingham, AL; †Department of Surgery, University of Illinois College of Medicine, Chicago, IL; ‡Accreditation Council for Graduate Medical Education, Chicago, IL; §Department of Surgery, Washington University in St Louis, St Louis, MO; ‖Department of Surgery, University of California, San Francisco, San Francisco, CA; and ¶Department of Surgery, University of Utah Health, Salt Lake City, UT.

✉ pzmijewski@uabmc.edu.

The authors report no conflicts of interest.

DOI: 10.1097/SLA.0000000000006373

Table 1. Demographics

	N = 6458 (%)
Non-URiM	5604 (86.8)
URiM	854 (13.2)
Male	3833 (59.4)
Female	2625 (40.6)

surgery residency programs that self-reported their case log data, illuminated disparities in case numbers. These studies represent the only published work to date examining case log disparities at a multi-institutional level. The ROPE Consortium found that underrepresented in medicine (URiM) residents performed 76 fewer cases than their non-URiM peers.[6] Likewise, female residents performed 37 fewer cases over 5 years when compared with their male peers.[7] A post hoc analysis revealed that residents in the lowest quartile for case volumes were more likely to be female and identify as URiM.[8] This work, though invaluable to our understanding of the United States training environment, represents data from <10% of all training programs accredited by the ACGME.

In this study, we expanded on the work of the ROPE Consortium through a review of the only national database of general surgery case logs in the United States. We analyzed case log reports submitted to the ACGME by chief residents applying for certification in the last 5 years (2017–present). Case logs were interrogated for total case numbers, as well as case type/complexity distribution by sex and race. This is the first time these data have been examined nationally.

METHODS

After institutional IRB approval, the Association of Graduate Medical Education case log database was queried. All case logs submitted by chief residents at ACGME-accredited programs were examined between the years of 2017 and 2022. Total major cases, cases by category, subcategory, residency program, sex of resident, and race of resident were obtained. Mean case numbers were compared using paired *t* tests. Statistical significance was defined as a *P* value of <0.05.

Male and female sexes (defined as the biological status of the resident, typically assigned at birth) were the reported variables in the ACGME database. Data on self-reported/identified sex of residents was not available for analysis.

The American Association of Medical Colleges' definition of URiM was adopted for the purposes of this study. Per the American Association of Medical Colleges, "URiM means those racial and ethnic populations that are underrepresented in the medical profession relative to their numbers in the general population."[9] African-American, Mexican-American, Native American (American Indian, Alaska Native, Native Hawaiian), and mainland Puerto Rican residents were considered URiM in this study due to present and historical national underrepresentation in the medical field.

RESULTS

Three hundred nineteen ACGME-accredited programs submitted chief resident case logs between 2017 and 2022. A total of 6458 residents logged cases. Of these, 3833 (59.4%) were males and 2625 (40.6%) were females. Eight hundred fifty-four (13.2%) were URiM residents, and 5604 (86.8%) were non-URiM residents (Table 1).

Mean major cases logged from 2017 to 2022 by all residents was 1113.45 (SD ± 160.65). Mean surgeon chief cases logged by all residents was 279.68 (SD ± 69.71). Mean teaching assistant cases logged by all residents was 43.05 (SD ± 21.62).

Males logged 1128.56 (SD ± 168.32) major cases, and females logged 1091.38 (SD ± 145.98; difference = 37.18, $P < 0.0001$). Males logged 282.71 (SD ± 72.59) surgeon chief cases, and females logged 275.25 (SD ± 65.03; difference = 7.46, $P < 0.0001$). Males logged 44.68 (SD ± 22.80) teaching assistant cases, and females logged 40.66 (SD ± 19.53; difference = 4.02, $P < 0.0001$).

Non-URiM residents logged (1115.96 SD ± 160.53) major cases, and URiM residents logged (1096.95 SD ± 160.57; difference = 19.96, $P = 0.001$). Non- URiM residents logged (279.91 SD ± 68.96) surgeon chief cases and URiM residents logged (278.13 SD ± 74.43; difference = 1.78, $P = 0.486$). Non- URiM residents logged (43.16 SD ± 21.40) teaching assistant cases, and URiM residents logged 42.31 (SD ± 23.02; difference = 0.85, $P = 0.286$; Table 2). Differences in total major case numbers by individual study year can be seen in Table 3.

Major case categories were likewise examined. Female residents logged significantly fewer cases in abdominal (378.96 male vs 359.93 female), alimentary tract (288.55 vs 285.10), endoscopy (135.57 vs 133.36), head and neck (65.77 vs 64.44), basic laparoscopy (216.37 vs 202.54), complex laparoscopy (137.55 vs 130.61), nonoperative trauma (69.28 vs 68.22), operative trauma (32.55 vs 31.53), thoracic (44.60 vs 40.53), and vascular (120.74 vs 114.61; $P < 0.05$) categories. Female residents logged significantly more breast cases than male residents (65.45 male vs 72.27 female; $P < 0.0001$). There was no statistically significant difference noted in the number of cases in the endocrine (33.22 vs 35), nonoperative trauma (68.87 vs 69.28), pediatric (39.86 vs 41.68), plastics (28.39 vs 27.81), skin and soft tissue (67.06 vs 70.68), and surgical critical care (49.25 vs 51.56) categories ($P > 0.05$). The most significant disparities were noted in the abdominal (difference = 16.77), basic laparoscopy (11.73), and complex laparoscopy (10.53) categories.

URiM residents logged significantly fewer cases in abdominal (376.85 non-URiM vs 364.02 URM), alimentary

Table 2. Difference in Case Numbers by Type (2017–2022)

Case type	Non-URiM vs URiM case # difference	URiM more or less cases	*P*	Male vs female case # difference	Female more or less cases	*P*
Major cases	19.96	Less	0.001	37.18	Less	<0.0001
Teaching assistant	0.85	Less	0.286	4.02	Less	<0.0001
Surgeon chief	1.78	Less	0.486	7.46	Less	<0.0001

Copyright © 2024 Wolters Kluwer Health, Inc. All rights reserved.

Table 3. Difference in Major Case Numbers By Year

Year	Non-URiM vs URiM case # difference	URiM more or less cases	*P*	Male vs Female case # difference	Female more or less cases	*P*
2017–2018	37.32	Less	0.011	25.86	Less	0.004
2018–2019	16.40	Less	0.198	35.71	Less	<0.001
2019–2020	12.23	Less	0.344	43.08	Less	<0.001
2020– 2021	23.73	Less	0.075	34.70	Less	<0.001
2021–2022	5.72	Less	0.646	44.95	Less	<0.001

tract (301.68 vs 294.93), endoscopy (134.52 vs 127.14), basic laparoscopy (220.59 vs 211.91), and complex laparoscopy (149.18 vs 144.34) categories ($P < 0.05$). URiM residents logged more major cases in the operative trauma category (32.86 non-URiM vs 34.98 URiM, $P = 0.003$). No statistically significant difference was noted in the breast (69.05 non-URiM vs 67.90 URiM), endocrine (33.07 vs 32.42), head and neck (62.86 vs 63.58), nonoperative trauma (67.42 vs 66.76), pediatric (38.82 vs 38.41), plastics (29.18 vs 29.89), skin and soft tissue (68.74 vs 69.36), surgical critical care (52.50 vs 54.55), thoracic (42.50 vs 42.20), and vascular (117.30 vs 117.94) categories ($P > 0.05$; Table 4).

Case subcategories were likewise examined. Female residents logged fewer cases in the Access (28.54 male vs 27.15 female), anastomosis, repair, exposure or endarterectomy (37.04 vs 34.05), appendix (78.44 vs 75.46), biliary (145.02 vs 136.66), colonoscopy (70.96 vs 67.63), esophagus (14.94 vs 13.84), hernia (142.52 vs 135.87), large intestine (74.82 vs 71.83), liver (14.78 vs 13.93), open thoracotomy (11.44 vs 10.66), pancreas (12.15 vs 11.53), stomach (48.25 vs 46.16), and upper endoscopy (56.05 vs 54.11) subcategories ($P < 0.05$). Females logged more axilla (16.03 male vs 17.62 female) and mastectomy (41.85 vs 44.18) cases ($P < 0.05$). No statistically significant difference was found in the Anorectal (39.77 male vs 40.26 female), small intestine (48.43 vs 47.61), team leader resuscitation (16.37 vs 16.46), and thyroid/ parathyroid (28.95 vs 29.74) subcategories ($P > 0.05$).

URiM residents logged fewer cases in the access (28.14 non-URiM vs 26.84 URiM), anorectal (40.17 vs 38.71), biliary (142.33 vs 137.01), colonoscopy (70.04 vs 66.73), hernia (140.63 vs 134.51), large intestine (73.87 vs 71.89), and upper endoscopy (55.70 vs 52.41) subcategories ($P < 0.05$). No difference in case numbers was found in the anastomosis, repair, exposure, or endarterectomy (35.81 non-URiM vs 35.94 URiM), appendix (77.42 vs 75.93), axilla (16.69 vs 16.54), esophagus (14.56 vs 14.02), liver (14.43 vs 14.48), mastectomy (42.91 vs 42.03), open thoracotomy (11.13 vs 11.10), pancreas (11.91 vs 11.79), small intestine (48.15 vs 47.70), stomach (47.51 vs 46.67), team leader resuscitation (16.26 vs 17.34), and thyroid/ parathyroid (29.38 vs 28.54) subcategories ($P > 0.05$; Table 5).

DISCUSSION

Structural disparities in residency training have been shown to disproportionally affect residents who are URiM and female. In the case of URiM residents, these disparities can come in the forms of macro and micro aggression, bias, discrimination, and denial of equal learning and engagement opportunities. URiM residents have described suffering from lower expectations, social isolation, and harsher punishment for similar mistakes as their peers.[10–12] In a national survey study, Black residents stated they did not feel that they "fit in" at their training programs, could not count on their peers for help, and were more likely to feel that they would require additional post-residency training to be practice ready.[13]

Sex disparities in surgical training have been demonstrated in a number of studies detailing differences in operative autonomy,[14–17] awards distribution,[18] weaker feedback,[19] poor letters of recommendation,[15,20–22] and varying rates of personal experiences with sex-based

Table 4. Difference in Case Numbers by Category 2017-2022

Case category	Non-URiM vs URiM case # difference	URiM more or less cases	*P*	Male vs Female case # difference	Female more or less cases	*P*
Abdominal	12.83	Less	<0.0001	16.77	Less	<0.0001
Alimentary tract	6.75	Less	0.005	9.49	Less	<0.0001
Breast	1.15	Less	0.269	5.75	More	<0.0001
Endocrine	0.65	Less	0.383	0.67	More	0.188
Endoscopy	7.38	Less	0.000	5.57	Less	<0.0001
Head and neck	0.72	More	0.433	2.37	Less	<0.0001
Laparoscopic basic	8.68	Less	<0.0001	11.73	Less	<0.0001
Laparoscopic complex	4.84	Less	0.008	10.53	Less	<0.0001
Nonoperative trauma	0.66	Less	0.716	2.77	Less	0.028
Operative trauma	2.12	More	0.003	1.62	Less	0.001
Pediatric	0.41	Less	0.552	0.62	More	0.186
Plastics	0.71	More	0.408	1.07	Less	0.073
Skin and soft tissue	0.62	More	0.528	0.72	More	0.291
Surgical critical care	0.3	Less	0.096	0.5	More	0.500
Thoracic	0.3	Less	0.761	3.64	Less	<0.0001
Vascular	0.64	More	0.765	9.22	Less	<0.0001

Table 5. Difference in Case Number by Subcategory 2017-2022

Case subcategory	Non-URiM vs URiM case # difference	URiM more or less cases	*P*	Male vs Female case # difference	Female more or less cases	*P*
Access	1.3	Less	0.048	0.82	Less	0.005
Anastomosis, repair, exposure, or endarterectomy	0.13	More	0.880	2.99	Less	2.99
Anorectal	1.46	Less	0.019	0.29	More	0.291
Appendix	1.49	Less	0.128	2.98	Less	<0.0001
Axilla	0.15	Less	0.665	0.95	More	<0.0001
Biliary	5.32	Less	0.001	8.36	Less	0.001
Colonoscopy	3.31	Less	0.009	3.33	Less	<0.0001
Esophagus	0.54	Less	0.100	1.1	Less	<0.0001
Hernia	3.73	Less	<0.0001	6.65	Less	<0.0001
Large intestine	1.98	Less	0.026	2.99	Less	<0.0001
Liver	0.05	More	0.873	0.85	Less	<0.0001
Mastectomy	0.88	Less	0.258	2.33	More	<0.0001
Open thoracotomy	0.03	Less	0.902	0.78	Less	<0.0001
Pancreas	0.12	Less	0.637	0.62	Less	0.002
Small intestine	0.82	Less	0.050	0.45	Less	0.420
Stomach	0.84	Less	0.427	2.09	Less	0.008
Team leader resuscitation	1.08	More	0.063	0.09	More	0.834
Thyroid or parathyroid	0.84	Less	0.245	0.79	More	0.143
Upper endoscopy	3.29	Less	<0.0001	1.94	Less	0.004

discrimination.[2,23–26] Female residents may have more limited access to quality mentorship[27] and professional development opportunities,[28] which can lead to pervasive deleterious patterns throughout their career development, for example, inequitable compensation.[29] Multiple studies have shown that female residents express less satisfaction with their training experience.[30–32]

Our study represents the first of its kind, in that it examines the case logs of all accredited ACGME programs and provides objective evidence of training disparities demonstrated in the previously mentioned subjective studies. Our results quantify the feelings of generations of females and URiM residents who have experienced overt or, more commonly, covert discrimination. Here, we have demonstrated that URiM residents and female residents log significantly fewer major cases than their male and non-URiM peers. When looking at surgeon chief and teaching assistant cases, there was no significant difference noted between cases submitted by URiM versus non-URiM residents. However, male residents reported significantly more in both categories than their female peers.

Interestingly, the gap between URiM and non-URiM residents was much smaller than that of the case discrepancy found in the ROPE consortium study[6] (19.96 fewer cases vs 76), demonstrating that there is likely a large amount of heterogeneity between programs. Analysis of disparities by geographic region, as well as program type (academic, community, hybrid, or military), would be a useful adjunct in understanding trends in disparities between programs. This also supports the need for programs to do their own individual assessments of case numbers to ensure no disparities in operative experience among their own trainees.

The gap in submitted cases between male and female residents was pronounced and is consistent with the findings in the ROPE study (37 vs 37.18 fewer cases). This gap was consistent and statistically significant throughout the entire study period and in most case categories (Table 3). A difference of 30 to 40 cases is equivalent to 1 to 3 months of surgical training and is a concerning national trend deserving the attention of every training program and our governing institutions.

The gap between URiM and non-URiM residents did seem to narrow over time, whereas the gap between female and male residents did not. This finding was a surprising outcome of our study, and merits more investigation, particularly in the qualitative realm, to determine why these disparities persist for female residents.

Our study is not without significant limitations. We are limited in our conclusions, as case logs are merely a reflection of what residents report, rather than the reality of what they perform. In addition, some residents may have lower case numbers due to taking time for parental leave, and this variable is not possible to account for in our data set. Our data set is also limited by our inability to measure program-level effects, as well as variability between. Individual programs introduce an additional source of bias.

Lastly, it is important to note that in no way are the authors of this study implying that URiM residents or female residents are poorly trained relative to their male and non-URiM peers. Recent initiatives by the ABS to move towards a more Competency-Based Medical Education model based on operative entrustment[33–37] have made raw case numbers as a direct measure of competency a thing of the past. Case number disparities by race and sex can instead be considered a "canary in the coal mine," suggestive of insidious structural bias in surgical education that must be urgently confronted and corrected.

CONCLUSIONS

URiM and female residents log fewer major cases during general surgery residency training than their non-URiM and male peers. A difference of 30 to 40 cases is equivalent to 1 to 3 months of surgical training and is a concerning national trend deserving the attention of every training program and our governing institutions.

REFERENCES

1. Foley KE, Izquierdo KM, von Muchow MG, et al. Colon and rectal surgery robotic training programs: an evaluation of gender disparities. *Dis Colon Rectum*. 2020;63:974–979.
2. Gurgel RK, Cardon BR, Allen CM, et al. Evaluating gender parity in operative experience for otolaryngology residencies in the United States. *Laryngoscope*. 2020;130:1651–1656.
3. Gong D, Winn BJ, Beal CJ, et al. Gender differences in case volume among ophthalmology residents. *JAMA Ophthalmol*. 2019;137:1015–1020.
4. Reyes-Capo DP, Yannuzzi NA, Chan RVP, et al. Gender differences in self-reported procedural volume among vitreoretinal fellows. *Retina*. 2021;41:867–871.
5. Eruchalu CN, He K, Etheridge JC, et al. Gender and racial/ethnic disparities in operative volumes of graduating general surgery residents. *J Surg Res*. 2022;279:104–112.
6. Eruchalu CN, Etheridge JC, Hammaker AC, et al. Racial and ethnic disparities in operative experience among general surgery residents. *Ann Surg*. 2024;279:172–179.
7. Winer LK, Kader S, Abelson JS, et al. Disparities in the operative experience between female and male general surgery residents. *Ann Surg*. 2023;278:1–7.
8. Price AD, Foote DC, Woeste MR, et al. Defining the disparity: a multi-institutional analysis of factors associated with decreased resident operative experience. *J Surg Res*. 2024;293:647–655.
9. American Association of Medical College. Underrepresented in medicine definition. Accessed June 14, 2024. https://www.aamc.org/what-we-do/equity-diversity-inclusion/underrepresented-in-medicine
10. Liebschutz JM, Darko GO, Finley EP, et al. In the minority: black physicians in residency and their experiences. *J Natl Med Assoc*. 2006;98:1441–1448.
11. Khubchandani JA, Atkinson RB, Ortega G, et al. Perceived discrimination among surgical residents at academic medical centers. *J Surg Res*. 2022;272:79–87.
12. Yuce TK, Turner PL, Glass C, et al. National evaluation of racial/ethnic discrimination in us surgical residency programs. *JAMA Surg*. 2020;155:526–528.
13. Wong RL, Sullivan MC, Yeo HL, et al. Race and surgical residency: results from a national survey of 4339 US general surgery residents. *Ann Surg*. 2013;257:782–787.
14. Meyerson SL, Odell DD, Zwischenberger JB, et al. The effect of gender on operative autonomy in general surgery residents. *Surgery*. 2019;166:738–743.
15. Al Jabri A, Bhat H, Abelson JS, et al. The presence of implicit gender bias in colon and rectal surgery residency letters of recommendation. *Dis Colon Rectum*. 2023;66:848–856.
16. Chen JX, Chang EH, Deng F, et al. Autonomy in the operating room: a multicenter study of gender disparities during surgical training. *J Grad Med Educ*. 2021;13:666–672.
17. Ofshteyn A, Steinhagen E. Surgical education: disparities in education may impact the quality and likelihood of completion of training. *Clin Colon Rectal Surg*. 2023;36:315–320.
18. Kuo LE, Lyu HG, Jarman MP, et al. Gender disparity in awards in general surgery residency programs. *JAMA Surg*. 2021;156:60–66.
19. Gerull KM, Loe M, Seiler K, et al. Assessing gender bias in qualitative evaluations of surgical residents. *Am J Surg*. 2019;217:306–313.
20. Oslock WM, Lansing SS, Coleman LR, et al. Gender bias in colorectal surgery fellowship letters of recommendation. *Am J Surg*. 2024;227:198–203.
21. Turrentine FE, Dreisbach CN, St Ivany AR, et al. Influence of gender on surgical residency applicants' recommendation letters. *J Am Coll Surg*. 2019;228:356–365e3.
22. Hoffman A, Grant W, McCormick M, et al. Gendered differences in letters of recommendation for transplant surgery fellowship applicants. *J Surg Educ*. 2019;76:427–432.
23. Bruce AN, Battista A, Plankey MW, et al. Perceptions of gender-based discrimination during surgical training and practice. *Med Educ Online*. 2015;20:25923.
24. Schlick CJR, Ellis RJ, Etkin CD, et al. Experiences of gender discrimination and sexual harassment among residents in general surgery programs across the US. *JAMA Surg*. 2021;156:942.
25. Barnes KL, Dunivan G, Sussman AL, et al. Behind the mask: an exploratory assessment of female surgeons' experiences of gender bias. *Acad Med*. 2020;95:1529–1538.
26. Baker S, Gleason F, Lovasik B, et al. Relationship between burnout and mistreatment: who plays a role? *Am J Surg*. 2021;222:1060–1065.
27. Rakestraw SL, Chen H, Corey B, et al. Closing the gap: Increasing female representation in surgical leadership. *Am J Surg*. 2022;223:273–275.
28. Dream S, Woolfolk M, Chen H. Gender role incongruency in general surgery applicants. *Am J Surg*. 2022;224:900–902.
29. Morris M, Chen H, Heslin MJ, et al. A structured compensation plan improves but does not erase the sex pay gap in surgery. *Ann Surg*. 2018;268:442–448.
30. Mullins CH, Lindeman B. Ending the burnout gender gap in surgery. *Am J Surg*. 2022;223:608.
31. Yeo H, Viola K, Berg D, et al. Attitudes, training experiences, and professional expectations of us general surgery residents. *JAMA*. 2009;302:1301.
32. Entezami P, Franzblau LE, Chung KC. Mentorship in surgical training: a systematic review. *Hand*. 2012;7:30–36.
33. Lindeman BM, Petrusa ME Entrustable Professional Activities (EPAs) and Applications to Surgical Training. Accessed February 11, 2024. https://www.facs.org/education/division-of-education/publications/rise/articles/entrustable
34. Lindeman B, Brasel K, Minter RM, et al. A phased approach: the general surgery experience adopting entrustable professional activities in the United States. *Acad Med*. 2021;96:S9–S13.
35. Brazelle M, Zmijewski P, McLeod C, et al. Concurrent validity evidence for entrustable professional activities in general surgery residents. *J Am Coll Surg*. 2022;234:938–946.
36. Brasel KJ, Lindeman B, Jones A, et al. Implementation of entrustable professional activities in general surgery. *Ann Surg*. 2023;278:578–586.
37. Montgomery KB, Mellinger JD, Jones A, et al. Validity of entrustable professional activities in a national sample of general surgery residency programs. *J Am Coll Surg*. 2024;238:376–384.

DISCUSSANT

Dr Steven Stain (Burlington, MA)

Thank you, Dr Fazendin. Dr Brasel.

Dr Karen Brasel (Portland, OR)

This well-presented study examined traditional and quantifiable metrics of surgical training, the case log. The authors have done a thorough job of looking at it through an equity lens. Do our female residents and those who are URiM have an equal training opportunity? As we consider competency-based education foreshadowing the upcoming presidential panel, understanding these findings and how they relate to competence and the opportunity to both gain and demonstrate competence in our trainees will be incredibly important.

I have so many questions, but I am a rule follower, so I will limit myself to 2. My first question is whether this is an effect that is playing out at the individual level or the program level. This is highlighted by the differences in the case-logging differences in your study compared with the ROPE study and perhaps more dramatically by the fact that the differences that you found varied by case type. There were some cases, actually, where female residents and URiM residents had more cases than the majority of residents. We know that URiM residents and female

residents are not equally distributed across training programs, and we also know that case type and average case numbers are not equally distributed. To use clinical trial language, is there a center effect? What is the difference in program makeup in terms of percent female and URiM, and does this affect your findings?

My second question has to do with the difference between chief cases and major cases. Can you speculate as to why there is a difference in major cases but not chief cases for URiM residents, but the difference remains for female residents? You highlight in your manuscript the importance of a chief year. If there is no difference in the chief year for URiM residents, as well as a decline in the difference over time, is what you are describing a distinction without a true difference? Although we might argue over the importance of absolute numbers as a measure of surgical skill or competence, there is no question that they can be a proxy for access and opportunity. For our patients and for the future of surgery, we need to ensure that access is equitable.

Thank you again for your work, and I look forward to further discussion.

Response From Jessica Fazendin

Dr Brasel, thank you so much. To our president and moderators, thank you for this invitation to discuss Dr Brasel's and the audience's questions.

To start off with, first to your question, is this an individual issue versus a programmatic effect, just to go and highlight for those in the audience that may not be as familiar with the ROPE study, was a study encompassing 21 major quaternary academic referral centers representing $<10\%$ of all ACGME surgical training programs. They demonstrated a difference in 76 major cases in those URiM versus our study representing a difference of 20. We think that this most likely reflects the heterogeneity seen in training programs across the country, both geographically as well as by program type. In our study, all programs, including academic, community, hybrid, and even military programs, are represented. There is more geographic diversity in our study, obviously encompassing every program within the nation as compared with the ROPE consortium.

To address the issue of programmatic effect, use 2 examples that were not demonstrated in the presentation but would be hypotheticals. Are underrepresented individuals in medicine placed more at those academic programs or in urban centers that may perform cases like operative trauma as compared with a community or rural program? Second, are more women distributed in programs that may be in the Northeast or have lower total operative cases by program alone? Unfortunately, our methodology was not centered around answering that question, though it is well worth exploring in the future.

To answer your second question, the only statistically significant difference in URiM versus non-URiM groups was major cases, but not chief cases. This leads us to believe that underrepresented individuals or marginalized groups may be more vulnerable in their early years. As I am sure many people can understand, the graduated autonomy in advancing years of residency also likely leads to more responsibility in scheduling cases. This may represent a "catch-up" effect because they can assign themselves those cases. That is one possibility.

We were pleased to see that the statistical significance did decrease with the years represented in this study. However, there still is a numerical difference, and as the presidential address yesterday and especially Dr Clarke's presentation just before this session addressed, there still are problems. We can clearly see in our data that though there are still numerical differences with underrepresented individuals, there still is a statistically significant difference in female versus male residents throughout the entire study period.

Thank you, Dr Brasel, for your questions.

Dr Stephen Wigmore (Edinburgh, Scotland)

I am sorry if this is a dumb question and I am not familiar with your system, but what responsibilities do residency program directors bear for making sure that the playing field is level? Presumably, there is an annual review process to make sure that residents are making the appropriate progress that they should.

Dr Jessica Fazedin (Birmingham, AL)

So speaking not as a program director, I know that being a clinical competency committee member at my local program, we do review residents twice a year and look at their case minimums. However, I do not believe or have not heard of anything in the literature or at the Association of Program Directors in Surgery that individual program directors are looking at that specifically, but our results do encourage others to look at their programs individually.

Thank you.

Dr John Potts (Chicago, IL)

John Potts, Chicago. I really appreciate the authors bringing this paper to American Surgical. I think it is a very important topic with important implications. I do think that the authors failed to mention, at least in the slides, one very important limitation of the study, which is these are self-reported data.

Until we have either qualitative studies done or we have implementation of a system in which every case is reported by the attending and the resident in cooperation with each other, we will never really know how many of these cases are surgeon chief, surgeon junior, teaching assistant, or assistant cases. We will never know that about any group in any setting. Again, I really appreciate you bringing this forward, but I think that is a huge limitation and one that we need to correct in the way that cases are reported by residents.

Response From Jessica Fazendin

Thank you so much for that point. I will say with regard to the logging effect or a potential logging effect, 2 studies come to mind, one in 2016 and the other in 2022, where yes, self-reported data were analyzed by programs, and looking at the accuracy of case logging, they identified a wide gap between 76% and 95% accuracy in case logging by an individual. It does not lead me to think that there is a difference in URiM or female residents as compared with their counterparts, non-URiM or male residents, respectively. But it is a great point. When we did see the differences, especially in females, being statistically significant through all 5 years, this led us to perform a qualitative review, interviewing residents at almost 2 dozen programs. It is currently unpublished data. However, when we coded for qualitative themes, we found that some of the barriers to reaching the OR were increased administrative work on the

floor, both internal and external pressures, but again, that is self-reported data as well, so I appreciate your point.

Dr E. Christopher Ellison (Powell, OH)

Thank you. I enjoyed the paper very much. I would like to comment on your statistical analysis, so I wonder if your data set is normally distributed and if it is not, should you be using medians and interquartiles versus the analysis that you did?

As a second point, I would say that although it is significant differences, is it really substantial and significant in terms of the outcomes of the residents?

Thank you very much.

Response From Jessica Fazendin

Dr Ellison, thank you for your question. First of all, we did go back and forth within the group and did have some statisticians represented on our authorship on the paper perform both parametric and nonparametric tests to try to address that as well. We found similar results.

With regard to this being statistically versus clinically significant, depending on what service you're on or the program type that you are at, this may represent between 1 and 3 months of surgical training where, you know, as we move into competency-based education, this issue may go away. However, it does leave some room for inference as to why there actually is a difference of getting to the OR, OR experience, all of that, and hopefully we will explore that a little bit more through qualitative efforts, as well as the studies that are to come with competency-based education.

Dr Joseph Losee (Pittsburgh, PA)

Sorry if I missed these data and you presented it, but were you able to look at the URiM group as a set and then compare males and females to see whether there was intersectionality between race and sex?

Response From Jessica Fazendin:

I am so glad that you brought that point up. That is one of the things that we are actively working on, but we just did not have the data ready for this paper or presentation. When it comes to intersectionality, let us say a Black woman, that group is going to represent about 5% or less of the total demographic studied, and so we would have to use inferential statistics, and we are exploring whether or not that is feasible.

Dr Joseph Losee (Pittsburgh, PA)

All right, thanks.

Dr Orlando Kirton (Abington, PA)

Orlando Kirton from Jefferson Abington Hospital. I too applaud American Surgical for having this paper on the program. In my program, the individuals who have the most challenges in attaining case numbers are African American women. It is not my African American men, and definitely, if you are female and you are a Black female, it becomes a double challenge, and it really requires true intentionality in terms of the residency program director, so there are both programmatic effects and individual effects, and again, I just really would like to hear from you because you seem fairly well versed on this topic, what are the solutions, what are the fixes?

Thank you.

Response From Jessica Fazendin

Thank you so much for that. I would say with big data, all it does is explore and expose an issue, and so it is our responsibility, I think, as people in this room to take this back to our home environments and our local regions and say, "This is something that is important we have to look at," and then our job is to really explore that intersectionality as much as we can, and I am a firm believer in that if somethings been studied in the literature, you should prove it one more time and then probably stop studying it. This is different. This is something that for as many people as can study this, publish it, and expose it, the better.

Dr Orlando Kirton (Abington, PA)

Thank you.

Dr Michael Brunt (St. Louis, MO)

Michael Brunt, St. Louis. Congratulations on bringing this to American Surgical. My question is did you look at the impact of flexibility in training potentially on case numbers? Maybe not every program has a lot of flexibility, but, for example, if you have a senior resident who is more interested in breast oncology and may be doing more of those cases, they can do 4, 5, or 6 of those cases in a day versus somebody who is in Hepato-Pancreato-Biliary that may only do 1 or 2, and could that have potentially impacted the numbers? It may not just be about numbers. It may be about the experience.

Response From Jessica Fazendin

Also a great point, and to go ahead and say that individual factors do not make a play in this would be false. I pursued endocrine surgery, so I was more inclined to pursue those cases during residency. Why that factor is, whether it is individual or external pressures on the individual trainee, we do not know, and we have not explored that as to what specialty the individual has chosen to pursue and then the case logs during their ACGME training.

Dr Steven Stain (Burlington, MA)

Thank you very much.