The American Journal of Surgery 221 (2021) 515–520

Contents lists available at ScienceDirect

The American Journal of Surgery

journal homepage: www.americanjournalofsurgery.com

Original Research Article

Gender perception bias of operative autonomy evaluations among residents and faculty in general surgery training

Caitlyn Cookenmaster [a], Saad Shebrain [a, *], Duncan Vos [b], Gitonga Munene [a], Lisa Miller [a], Robert Sawyer [a]

[a] Department of Surgery, Western Michigan University Homer Stryker M.D. School of Medicine, USA
[b] Department of Epidemiology and Biostatistics, Western Michigan University Homer Stryker M.D. School of Medicine, USA

ARTICLE INFO

Article history:
Received 17 July 2020
Received in revised form
21 October 2020
Accepted 5 November 2020

Keywords:
Zwisch scale
Operative autonomy
Intraoperative evaluation

ABSTRACT

Background: Resident operative autonomy (ROA) is critical and a shared responsibility of both faculty and residents during training. We hypothesize that there is a perception of gender bias in residents' performance as evaluated by faculty and residents.
Method: Over a period of five academic years, between July 2014 and June 2019, ROA was evaluated using the Zwisch score. Reciprocal evaluations were completed by faculty and residents.
Results: 39 surgeons (30 males, 9 females) and 42 residents (25 males, 15 females) completed 2360 evaluations (1180 by faculty, and a matched number by residents). PGY level was significantly associated with granting a higher level of autonomy. Gender of residents didn't affect the level of granted autonomy as evaluated by faculty. However, on self-evaluations, female residents rated their degree of autonomy lower than that of their male counterparts.
Conclusion: Gender did not influence the perception of autonomy granted as evaluated by faculty. However, on self-evaluations, female residents reported a lower degree of autonomy received.

Introduction

Autonomy, as defined in the dictionary, is "self-directing freedom and especially moral independence."[1] This definition perfectly describes one of the most critical principles in the profession of medicine and surgery. Operative autonomy implies two components: sound medical and surgical knowledge and procedural skills driven by a sound decision-making process. Assessing operative autonomy, especially in a way that considers gender equity, is challenging. There are many factors that can affect resident autonomy including resident clinical skill, attending confidence and experience, and case complexity. Recently, it has become more apparent that socioeconomic and cultural factors such as gender, medical legal climate, and productivity demands can also affect the level of autonomy.[2]

There are conflicting studies about the real or perceived degree of operative autonomy granted to female residents by attendings when compared to their male counterparts.[3] Some studies have shown female residents receiving lower scores when compared to their male colleagues,[4] while other studies have shown no significant impact of gender on autonomy scores.[5] It is, however, clear that female residents report higher incidences of discrimination and inappropriate workplace experiences than their male counterparts. Even though incidences of overt bias are relatively uncommon, the fact that, despite increasing numbers of women in surgery, there is underrepresentation of women in leadership roles suggests that implicit gender bias may still be driving some of these discrepancies.[6]

The Zwisch scale is a validated four-level scale used to grade the degree of autonomy and guidance the attending surgeon provides to the trainee during operative procedures.[3] This study aims primarily to evaluate the possible impact of resident gender on the resident and faculty perceived level of operative autonomy (outcome variable) as assessed by faculty on the Zwisch scale. The secondary aim was to evaluate the effect of postgraduate level, case complexity, years of faculty experience, and resident status (categorical vs. preliminary).

* Corresponding author. Western Michigan University Homer Stryker MD School of Medicine, 1000 Oakland Drive, Kalamazoo, MI, 49008, USA.
E-mail address: saad.shebrain@med.wmich.edu (S. Shebrain).

https://doi.org/10.1016/j.amjsurg.2020.11.016
0002-9610/

Methods

Study Design: This institutional IRB-approved retrospective study includes 2360 intraoperative evaluations of surgical residents by both faculty (1180) and residents (1180) over a period of five academic years, between July 2014 and June 2019.

Participant/setting: 39 surgeons (30 males, 9 females) completed evaluations of 42 residents in a single general surgery residency program during a 5-year period. Residents include: 26 categorical residents (18 male, 8 female) and 16 preliminary residents (9 male, 7 female). Residents were distributed as follows: 24 PGY1, 24 PGY2, 16 PGY3, 14 PGY4, and 13 PGY5. Table 1 includes resident and faculty characteristics, number of participants and evaluations completed, and description of cases evaluated. In 2014, we adopted the Global Rating Scale of Operative Performance evaluation and combined it with the Zwisch scale to create a modified evaluation form.[7]

Operative case coverage is assigned by senior residents. Evaluations for selected cases were assigned on a weekly basis by the Program Director to both faculty and residents. Both the attendings and residents received an e-mail notification via our online evaluation system, New Innovations®, alerting them to the pending evaluation. Every effort was made to ensure evaluations encompassed a diverse, level-appropriate variety of cases assigned to residents. Resident autonomy was scored by faculty and residents using the Zwisch scale. Evaluations were completed within three days of the case.

Statistical analysis and data collection

During the study period, 2927 evaluations were assigned to each faculty/resident pair. Faculty completed 1646 (56.24%) evaluations, and residents completed 2145 (73.28%). Of the total completed evaluations (3791), both faculty and residents each completed 1216 evaluations for same procedure, totaling 2432 evaluations (64.15%). After 36 evaluations were excluded due to missing data, the analyzed set contained 1180 evaluations completed by both residents and faculty for same procedure, for a total of 2360 evaluations. For the purpose of this study, we focused on the Zwisch autonomy scale component. Data were deidentified and entered into an Excel spreadsheet before analysis.

Frequency (percent) was reported for each dependent and independent variable. Independent variables include faculty and resident gender, resident-faculty gender combination, complexity of cases, postgraduate level, faculty experience (<10 years vs. 10+ years), and resident status (categorical vs. preliminary). The dependent variable of resident autonomy was measured using Zwisch scores.

A cumulative odds ordinal logistic regression with proportional odds was run to determine the effect of resident gender, faculty

Table 1
Demographic Characteristics of Participants (Faculty and Residents), Number of Residents in each PGY level, Evaluations Completed by Faculty, and ACGME Defined-Category Cases Completed by Residents.

Characteristics	Male n (%)	Female n (%)	Total N (%)
FACULTY	30 (77)	9 (23)	39 (100)
Specialty			
-General Surgery	10 (33.3)	2 (22.2)	12 (31)
-General Surgery Plus Subspecialty (GS + Trauma, Vascular, Bariatric)	6 (20)	3 (33.3)	9 (23)
-Subspecialty (Breast, Vascular, HBP, Trauma, Pediatrics, H&N, colorectal)	14 (46.7)	4 (44.4)	18 (46)
RESIDENTS	**27 (64.3)**	**15 (35.7)**	**42 (100)**
-Categorical	18 (66.7)	8 (53.3)	26 (62)
-Preliminary	9 (33.3)	7 (46.7)	16 (38)
Number of Residents at each PGY Level during Study Period[a]			
-PGY1, n = 24	17 (71)	7 (29)	24 (26.4)
-PGY2, n = 24	13 (54.2)	11 (45.8)	24 (26.4)
-PGY3, n = 16	11 (68.8)	5 (31.2)	16 (17.5)
-PGY4, n = 14	11 (78.6)	3 (21.4)	14 (15.4)
-PGY5, n = 13	10 (77)	3 (23)	13 (14.3)
Distribution of Evaluations Completed by Faculty at PGY Level			
-Cases Evaluated by Faculty	**1049 (88.9)**	**131 (11.1)**	**1180 (100)**
-PGY1	182 (17.3)	43 (32.8)	225 (19.1)
-PGY2	276 (26.3)	23 (17.6)	299 (25.3)
-PGY3	215 (20.5)	27 (20.6)	242 (20.5)
-PGY4	178 (17)	13 (9.9)	191 (16.2)
-PGY5	198 (18.9)	25 (19.1)	223 (19)
ACGME Defined-Category Cases Performed by Residents	**811 (68.7)**	**369 (31.3)**	**1180 (100)**
- Skin-Soft Tissue	70 (8.6)	44 (11.9)	114 (9.7)
- Breast	22 (2.7)	20 (5.4)	42 (3.6)
- Alimentary Tract	128 (15.8)	49 (13.3)	177 (15.0)
- Abdominal	405 (49.9)	176 (47.7)	581 (49.2)
- Vascular	67 (8.3)	24 (6.5)	91 (7.7)
- Endocrine	47 (5.8)	19 (5.1)	66 (5.6)
- Pediatric Surgery	11 (1.4)	10 (2.7)	21 (1.8)
- Thoracic Surgery	13 (1.6)	5 (1.4)	18 (1.5)
- Endoscopy	12 (1.5)	7 (1.9)	19 (1.6)
- Laparoscopic Complex	19 (2.3)	8 (2.2)	27 (2.3)
- Miscellaneous	17 (2.1)	7 (1.9)	24 (2.1)

*GS, General Surgery.
[a] During this study period (5 academic years), residents were analyzed according to their PGY level at the time of procedure and evaluations.

gender, complexity of cases, PGY, faculty years of experience, and resident status on autonomy score as rated by surgeon. The ordinal logistic regression model was used because the Zwisch scale is an ordinal variable and performing tests for difference in odds of achieving higher autonomy rating by gender will provide insight into whether there is an association between gender and autonomy rating. Each proportional odds model accounted for correlation due to repeated resident-faculty combinations. Unadjusted effects were assessed by including each independent variable alone in the model, and significant independent variables were included together to obtain a full model and adjusted effects. The proportional odds assumption was evaluated by running a logistic regression model for each possible dichotomized split of the ordinal autonomy outcome. SAS v9.4 was used for analysis. An alpha level of 0.05 was used.

Results

The median (interquartile range) of cases performed and evaluated by female and male residents was 17 (10, 32) cases and 18 (12, 49) cases, respectively. Female and male faculty evaluated a median (interquartile range) of 10 (4, 24) cases and 24 (11, 39) cases, respectively.

A cumulative odds ordinal logistic regression with proportional odds was run to determine the effect of resident gender, faculty gender, case complexity, postgraduate levels, faculty years of experience, and resident status on the higher level of autonomy as granted by faculty. The odd ratio (OR) of having a higher autonomy for female residents vs. male is 0.98 (95% CI, 0.89 to 1.074), $p = 0.644$, for preliminary residents vs. categorical is 0.755 (95% CI, 0.653 to 0.870), $p < 0.001$, with female faculty vs. male is 0.792 (95% CI, 0.688 to 0.911), $p = 0.001$, with faculty with less than 10 years of experience vs. more than 10 years is 1.3 (95% CI, 1.3 to 2.01), $p < 0.001$, and when performing basic procedures vs. advanced or complex is 1.48 (95% CI, 1.1 to 2.0), $p = 0.008$.

The significant variables included in the final ordinal logistic regression model consisted of resident gender, surgeon gender, the resident-surgeon gender interaction, procedure complexity, PGY level, and faculty years of experience (Table 2). Male and female residents rated by female surgeons had smaller odds of achieving a higher autonomy rating than male residents rated by male surgeons. Complex cases have significantly less odds of achieving higher autonomy than cases with basic difficulty. Faculty with experience of 10 or more years are less likely to score a higher level of autonomy than are faculty with experience of less than 10 years. Categorical residents have two times greater odds of achieving a higher autonomy rating relative to preliminary residents.

Our results indicate no gender effect regarding the level of autonomy granted to the trainee by surgeon (Fig. 1A). From the perspective of the faculty, all residents of the same postgraduate level received equivalent progressive autonomy. The odds of achieving a higher autonomy score increase about two-fold relative to PGY1 with every step in PGY level. From the perspective of the residents, both male and female, the results are mixed. At PGY1, PGY2, and PGY5 levels, residents granted themselves equivalent progressive autonomy. However, there is an autonomy gap at PGY3 and PGY4, where male residents' scores are higher than female residents'. There was not a significant difference between male and female residents with regard to the distribution of ACGME defined category cases ($p = 0.105$) and case complexity ($p = 0.059$) (Fig. 2). We found a moderate correlation between resident operative autonomy as rated by attending and PGY level ($r = 0.497$, $p < 0.001$).

Discussion

The primary aim of this study was to determine whether resident gender impacts the perceived level of autonomy granted in the OR. We found that, in our study, gender did not impact the level of perceived autonomy at all PGY levels of training or when looking at the entire study cohort. Two recent, large retrospective studies evaluating the impact of resident gender on the degree of autonomy granted demonstrated conflicting results. The first study by Lane[5] included analysis of evaluations completed by faculty only, and the second study by Myerson[4] included analysis of a combination of evaluations completed by faculty only, residents only, and both faculty and residents. Our result is in accordance with the study by Lane et al.[5] that demonstrated resident gender did not influence the perceived level of meaningful operative autonomy as defined by achieving level 3, "passive help," or level 4, "no help," on the Zwisch scale. However, it is in contradiction to those reported in a study by Myerson et al.,[4] which showed that women receive less autonomy in the operating room than men. Our study included evaluations completed by both faculty and residents for the same procedures. This resulted in a more homogenous sample, therefore mitigating the potential selection bias of including evaluations from faculty only or residents only. There could be additional reasons for these differences including study institution-specific variables and size of the study population. The differences in the outcomes are not as important as evaluating specific programs' outcomes with the intention of identifying any potential bias and

Table 2
Final proportional odds model.

Independent variable	Level, n (%)	Odds Ratio	LCL	UCL	p-value
Resident-Faculty Gender Combination	F - F, 41 (3.5)	0.48	0.25	0.91	0.0251
	F - M, 328 (27.8)	1.08	0.73	1.60	0.7086
	M – F, 90 (7.6)	0.61	0.40	0.92	0.0193
	M – M, 721 (61.1) [ref]	-	-	-	-
Procedure Complexity	Advanced, 367 (31.1)	0.78	0.59	1.03	0.0833
	Complex, 476 (40.3)	0.66	0.48	0.91	0.0113
	Basic, 337 (28.6) [ref]	-	-	-	-
PGY Level	PGY2, 299 (25.3)	3.09	2.11	4.50	<0.0001
	PGY3, 242 (20.5)	5.20	3.36	8.06	<0.0001
	PGY4, 191 (16.2)	11.13	6.80	18.21	<0.0001
	PGY5, 223 (19)	23.07	13.75	38.71	<0.0001
	PGY1, 225 (19.1) [ref]	-	-	-	-
Faculty Experience	10+ years, 754 (64)	0.62	0.46	0.84	0.0019
	<10 years, 426 (36) [ref]	-	-	-	-
Resident Status	Preliminary, 154 (13)	0.49	0.32	0.75	0.001
	Categorical, 1026 (87) [ref]	-	-	-	-

F=Female, M = Male, ref = reference, PGY = postgraduate level, LCL = lower confidence level, UCL = upper confidence level.

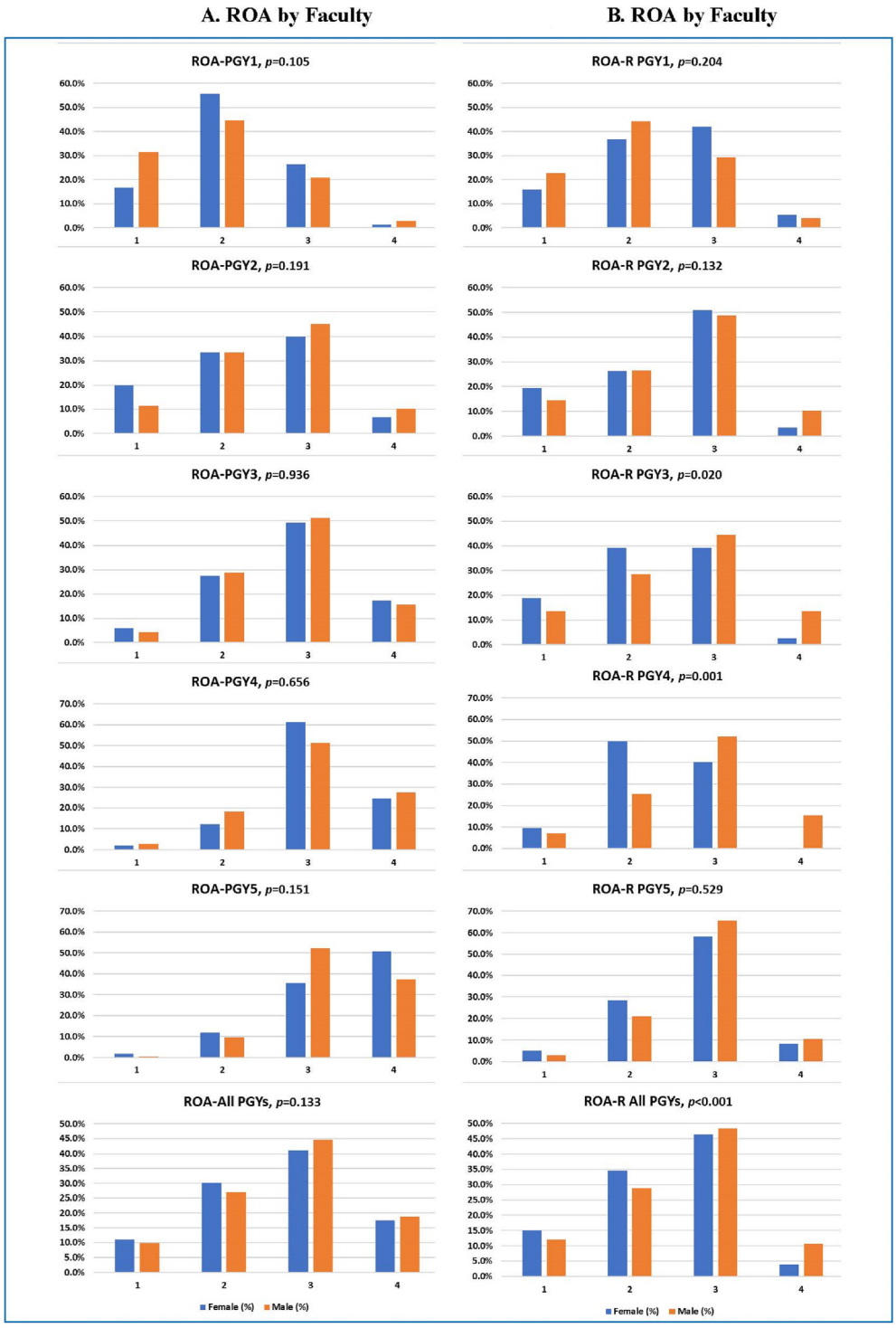

Fig. 1. Autonomy as graded by faculty (ROA) and residents (ROA-R) at Different PGY levels.

devising strategies to remedy some of these discrepancies. We believe these efforts should be undertaken at the institutional and program level given the differences that exist between programs.

When looking at resident self-reporting on perception of autonomy, we found that female residents as whole, and specifically as PGY3s and PGY4s, reported lower levels of perceived autonomy compared to their male counterparts. Previous studies have demonstrated that residents tend to underestimate their abilities when compared to their faculty assessment. This may not be a detrimental attribute because one may argue that this would produce a safer surgeon who is cognizant of the abilities or limitations.[8] However, if there are gender differences in the level of perceived autonomy, the impact of these differences in both learning and long-term self-confidence may be substantial and long standing. There may be plausible reasons for these gender differences in perception including the fact that men have been

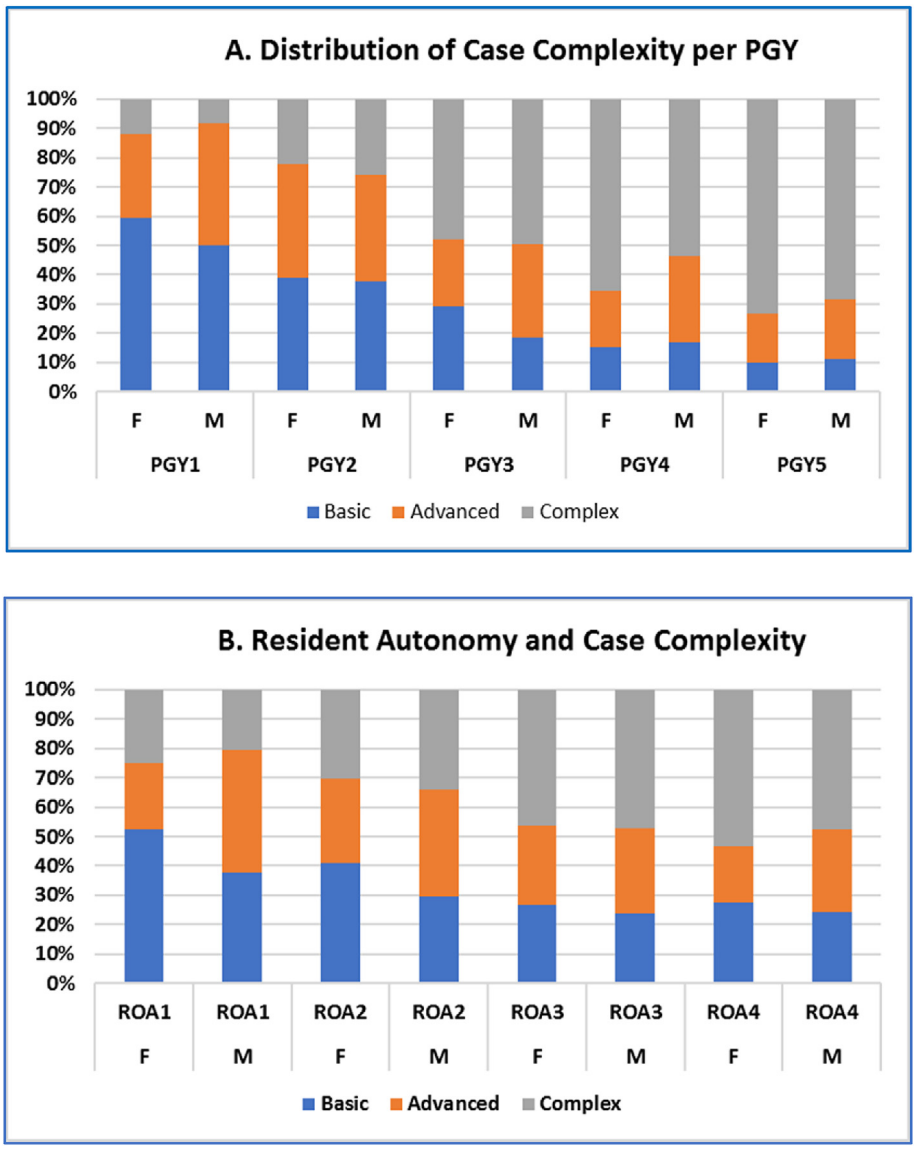

Fig. 2. Case complexity per PGY level, and autonomy granted.

reported to be more self-promoting than women.[9] It has also been reported that being assertive and competitive is perceived more favorably in men than similar behaviors in women.[10] Therefore, men may feel more empowered to ask for autonomy and, as a result, report higher levels of perceived autonomy. It is important that faculty and institutional leadership actively strive to create an environment where female residents can comfortably advocate for more autonomy without fear as a step to build the confidence the faculty surgeon has in their abilities.[4]

When we evaluated faculty factors that affected perceived level of autonomy, we found a weak but inverse correlation between the number of years the surgeon has been practicing and the level of autonomy granted. This relationship has been demonstrated in other studies.[5,7] One explanation may be that the younger surgeons operate with more senior residents and perform less complex procedures (Fig. 2). We also found that female faculty appeared to grant lower autonomy to both male and female residents compared to male faculty as reported by both residents and faculty. It was suggested that the level of autonomy granted by female faculty may not necessarily be lower than that for male faculty, but rather that female faculty's self-assessment of autonomy was, in itself, lower.[4] The reason for this difference is unclear and warrants further study especially given the lower numbers of female faculty included in our study.

Our study has limitations. First, the major limitation is the fact that autonomy was self-reported by residents and faculty, and not an independent observer. Second is the dependency on voluntary faculty participation. Many faculty do not complete all evaluations assigned. Third is the potential selection bias. Because all evaluations studied were completed for elective and non-emergent cases, the sample size and case diversity were reduced. Fourth, case complexity was assigned by the investigating surgeon and based on the average complexity of each procedure. This may not account for procedures that were more complex than typically experienced. Fifth, only matched evaluations (i.e., those completed by both faculty and residents for the same procedure) were included, which reduced the sample size of this study. Sixth, the results are from a single residency program with a limited number of participants (both in number of residents and number of female faculty), resulting in a relatively small sample. Although 23.1% (9/39) of faculty in our surgery department are female, only 11.1% of

evaluations (131) were completed by female faculty compared to 88.9% (1049) completed by male faculty.

Conclusion

Resident gender did not influence the perceived autonomy as evaluated by faculty. However, on self-evaluations, autonomy scores for female categorical residents are lower than their male counterparts. This difference in gender perception is an opportunity for programs to create a common ground in understanding the concept of consistent operative autonomy.

Competencies

Patient care, Medical knowledge.

References

1. https://www.merriam-webster.com/dictionary/autonomy.
2. Elfenbein DM. Confidence crisis among general surgery residents: a systematic review and qualitative discourse analysis. *JAMA Surg*. 2016;151(12):1166–1175. https://doi.org/10.1001/jamasurg.2016.2792.
3. DaRosa DA, Zwischenberger JB, Meyerson SL, et al. A theory-based model for teaching and assessing residents in the operating room. *J Surg Educ*. 2013;70(1):24–30. https://doi.org/10.1016/j.jsurg.2012.07.007.
4. Meyerson SL, Odell DD, Zwischenberger JB, et al. The effect of gender on operative autonomy in general surgery residents. *Surgery*. 2019;166(5):738–743. https://doi.org/10.1016/j.surg.2019.06.006.
5. Lane SM, Young KA, Hayek SA, et al. Meaningful autonomy in general surgery training: exploring for gender bias. *Am J Surg*. 2020;219(2):240–244. https://doi.org/10.1016/j.amjsurg.2019.11.035.
6. Garg Nupur. *Underrepresentation of Females in Academia: A Relative Rate Index and System Dynamics Model*vol. 1718. Yale Medicine Thesis Digital Library; 2012. http://elischolar.library.yale.edu/ymtdl/1718.
7. Shebrain S, Mahmood G, Munene G, Miller L, Collins J, Sawyer R. The role of manual dexterity and cognitive functioning in enhancing resident operative autonomy. *J Surg Educ*. 2019;76(6):e66–e76. https://doi.org/10.1016/j.jsurg.2019.06.002.
8. Minter RM, Gruppen LD, Napolitano KS, Gauger PG. Gender differences in the self-assessment of surgical residents. *Am J Surg*. 2005;189(6):647–650. https://doi.org/10.1016/j.amjsurg.2004.11.035.
9. Reuben E, Sapienza P, Zingales L. How stereotypes impair women's careers in science. *Proc Natl Acad Sci U S A*. 2014;111(12):4403–4408. https://doi.org/10.1073/pnas.1314788111.
10. Carnes M, Devine PG, Baier Manwell L, et al. The effect of an intervention to break the gender bias habit for faculty at one institution: a cluster randomized, controlled trial. *Acad Med*. 2015;90(2):221–230. https://doi.org/10.1097/ACM.0000000000000552.