Investigating Gender Differences in Faculty Evaluations by Trainees in a Gender-Balanced General Surgery Program

Jillian L. Angelo, MD, MS,[*,†] Ashkan Moazzez, MD, MPH,[*,†] Angela Neville, MD,[*,†] Christine Dauphine, MD,[*,†] Yazmin Lona[*,†] and Christian de Virgilio, MD[*,†]

[*]Department of Surgery, Harbor-UCLA Medical Center, Torrance, California; and [†]The Los Angeles Biomedical Research Institute, Los Angeles, California

PURPOSE: Women account for 21% of faculty positions in general surgery. In fields with lower female representation, female faculty receive lower evaluation scores by trainees compared to male faculty. At 42%, the female faculty representation in our general surgery department doubles the national average. We sought to determine if variations in faculty evaluations would be observed in a more gender-balanced general surgery program.

METHODS: Two years of faculty teaching evaluations by residents in a general surgery residency program were collected from the MedHub system. Total 3277 resident evaluations of 26 faculty members (11 female, 15 male) were analyzed. Seven areas (scored 1-7, with 1 = needs improvement and 7 = outstanding) were examined. Chi-square test was used to compare the percentage of male and female faculty members who scored a 6 or 7 in each category, and multivariate logistic regression analysis was used to determine the association of gender with the evaluation score, while adjusting for the number of encounters between the trainee and the faculty member.

RESULTS: There were no significant differences between male and female faculty in the "overall" evaluation score, nor in the "practice-based learning" and the "interpersonal and communication skills" categories. Female faculty had statistically significantly higher scores in "patient care", "professionalism," and "systems-based care" categories, whereas male faculty had higher evaluations in the "medical knowledge" category.

CONCLUSION: In a general surgery residency program with a relatively gender-balanced faculty, there was no gender difference in the "overall" evaluation of faculty by residents. However, there were gender differences in specific domains. These findings suggest that gender balance in teaching faculty may help eliminate previously observed teaching evaluation bias in the traditionally male dominated fields. (J Surg Ed 76:e132–e137. © 2019 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: General surgery, Faculty evaluations, Gender bias, Resident, ACGME core competencies

COMPETENCIES: Professionalism, Interpersonal and Communication Skills, Systems-Based Practice

INTRODUCTION

Teaching evaluations from both medical students and residents are a vital component of career advancement for faculty, as these evaluations affect retention, tenure, and promotion.[1–3] Prior research has shown that medical student evaluations are biased against female gender, particularly in surgery.[4] The reason for this bias is unclear; however, there appears to be an association with the proportions of women within a given specialty. In specialties with a high percentage of women faculty, evaluations of women tend to be better than those specialties with a low proportion of women faculty.[5] Surgery has historically had few female faculty, particularly in academic medicine, and there is even less female representation in the higher ranks of professorship. Although the number of women entering surgical careers and pursuing surgical academia are rising, the percentage of females at the full professor level remains

Presented at the APDS Meeting: Thursday, April 25 in Chicago, IL.
Correspondence: Inquiries to Christian de Virgilio, MD, Department of Surgery, Harbor-UCLA Medical Center, 1000 W. Carson Street, Torrance, CA 90502; e-mail: cdevirgilio@labiomed.org

  1931-7204/$30.00
https://doi.org/10.1016/j.jsurg.2019.06.007

low. Whether gender bias within evaluations may be contributing to the lack of female physicians in leadership roles in academia is unclear.

At our academic institution, women represent nearly half (43%) of our core general surgery residency teaching faculty, which is double the national average. We hypothesized that male and female faculty at our more gender-balanced institution would have similar overall evaluation scores. We further sought to determine whether there were gender differences in the evaluations with respect to the 6 Accreditation Council for Graduate Medical Education (ACGME) core competencies.

MATERIAL AND METHODS

This study was deemed exempt from human research by our institutional review board. Two years (from August, 2016 to September, 2018) of faculty teaching evaluations by residents in a general surgery residency program were collected from the MedHub system. Medhub is a web-based residency management system. Part of this system allows for evaluations of faculty by trainees. Evaluations of faculty by residents within the MedHub system are anonymous. The department chair, program director leadership, and program coordinators have access to these evaluations and know the identity of the faculty. In order to protect the identity of the faculty members, the evaluations were provided to the study team in a deidentified fashion by the program coordinator so the only identifier was the gender of the individual faculty (listed as "Male 1," "Female 1," "Male 2,", etc.). The identity of the residents completing the faculty evaluations is not recorded.

Inclusion criteria included all full-time faculty who were members of the teaching faculty within the general surgery residency program. Part-time faculty and full-time faculty that were not part of the general surgery residency training program were excluded. Our faculty includes 26 members. Eleven are female (42%) and 15 are male (58%). Our residency program consists of 6 categorical residents per year, 14 preliminary surgery residents at the PGY 1 level and 2 preliminary residents at the PGY 2 level. It is a 6-year program including 1 year of mandatory research. During the academic year 2016-2017, our residency consisted of 50 residents including 16 (32%) female residents and 34 (68%) male. During the academic year 2017-2018, the residency again was comprised of 50 total residents including 21 female (42%) and 29 male (58%). Our compliance rate for evaluations was 100% over those 2 years. There was no formal training on gender equality during the years the evaluations were collected.

The faculty evaluations assessed each of the 6 ACGME competencies (patient care, medical knowledge, practice based learning and improvement, interpersonal and communication skills, professionalism, and system based practice) as well as an Overall value to learning score (Table 1). Each category was rated from 1-7 (with 1 = needs improvement and 7 = outstanding) with a total of 19 sub-categories examined. The number of encounters between the faculty member and the trainee is listed by the trainee at the time of the evaluation and are scored from 1-3 (1 = 1-3 encounters, 2 = 4-9 encounters, and 3 = 9+ encounters).

Chi-square test was used to compare the percentage of male and female faculty members who scored a 6 or 7 in each category, and multivariate logistic regression analysis was used to determine the association of gender with the evaluation score, while adjusting for the number of encounters between the trainee and the faculty member. Statistical significance was determined as a p value of <0.05.

RESULTS

During the study period, there were a total of 3277 resident evaluations of the 26 faculty (Table 2). We reviewed a total of 1621 evaluations of female faculty (with range 13-333 evaluations per faculty member) and a total of 1656 evaluations of male faculty (with range 18-309 evaluations per faculty member). We adjusted for the number of encounters in order to minimize selection bias for those with very few (fewest = "Female 3" with 13 evaluations) or very many ("Female 7" with 333 evaluations) to minimize skewing the data based on these outliers.

There was no difference between male and female faculty in the "Overall Value to Learning" score. Within the 6 ACGME competencies, there were no significant differences in the "Practice Based-learning" or "Interpersonal and Communication Skills" scores. Female faculty had statistically significantly higher ratings in "Patient Care," with female faculty scoring higher in "Deeply Interested in Patient Care, Making Contributions to Their Management" (OR 1.5 95%CI[1.25-1.92], $p < 0.0001$), "Sensitive to Culture, Age, Gender, Disability" (OR 1.62 95%CI[1.32-1.97], $p < 0.0001$) and "Compassionate to Patient Needs When Performing Procedures" (OR 1.39 95%CI[1.12-1.73], $p = 0.003$). Female faculty scored significantly higher in 1 of the 4 subcategories of "Professionalism" "Keeps Appointments, Punctual, Does Not Leave Early, Attentive During Supervision" (OR 1.62 95%CI[1.31-2.00], $p < 0.001$) as well as 1 of the 2 subcategories of "System Based Practice" ("Advocates for Patients Within the Healthcare System," (OR 1.24 95%CI [1.01-1.53]),$p = 0.04$). These differences persisted after adjusting for number of encounters. In the "Medical Knowledge" category, male faculty scored significantly higher in "Appropriately Applies Knowledge of his/her

TABLE 1. Categories with Subcategories of Evaluations in the MedHub System

Category	Subcategory
Patient care	Deeply interested in patient care, making contributions to their management
	Sensitive to culture, age, gender, disability
	Compassionate to patient needs when performing procedures
Medical knowledge	Possesses excellent clinical acumen
	Displays investigatory and analytic thinking when discussing clinical problems
	Appropriately applies knowledge of his/her subject or speciality
Practice based learning and improvement	Demonstrates and fosters self-directed learning
	Ability to motivate, enthusiastic about subject
Interpersonal and communication skills	Presents material in an organized, clear manner; summarizes major points and provides emphasis
	Devotes appropriate amounts of time and discussion in an open and friendly manner
	Is approachable and receptive to questions
	Effectively teaches and demonstrates technical, surgical and procedural skills
Professionalism	Enjoys teaching and is enthusiastic about subject
	Keeps appointments, punctual, does not leave early, attentive during supervision
	Excellent role model as teacher and physician
	Commitment to the educational program
System based practice	Practices and teaches cost effective care
	Advocates for patients within the health care system
Overall	Overall value to your training

Subject or Specialty" when adjusted for number of encounters (OR 0.80 95%CI[0.64-0.99], p = 0.04). There was a trend towards significance for males in the "Possesses Excellent Clinical Acumen" after adjusting for number of encounters (OR 0.82 95%CI [0.67-1.01], p = 0.06).

DISCUSSION

This single institution, 2-year, retrospective study sought to determine whether there were gender differences in faculty evaluations by residents. Our program has a relatively high proportion of female faculty as compared to national averages. The study found that teaching evaluations were overall very positive, and there was no significant gender difference in evaluations with respect to "Overall value to training." However, within subcategories of the 6 ACGME competencies, women had significantly higher ratings in patient care, professionalism, and systems based practice and men had significantly higher ratings in medical knowledge. These findings indicate that male and female faculty are both highly valued and suggest that men and women may be perceived by residents as having different teaching strengths. These findings further support the notion that gender diversity enhances the overall teaching mission of a surgical training program.

Prior studies suggest faculty evaluations may be influenced by the gender of the professor, the gender of the student, and the field of study. At the undergraduate level, Basow found that ratings of male professors were unaffected by student gender, but female professors received the highest ratings from their female students and the lowest ratings from their male students.[6] Bachen et al. found through survey data that female students rated female faculty especially high and male faculty lower, whereas male students did not evaluate faculty differently based on gender.[7] Other studies suggest students tend to rate faculty of the same gender more favorably.[8] In a study of physician trainees, Fassiato et al. found that while there were no differences in faculty evaluations in subspecialties with higher female representation, female faculty evaluations were lower than male faculty in specialties with low female representation, such as those that were procedural in nature.[5] With a female representation of 21%, general surgery is considered to have low female representation. In contrast, our general surgery program is more gender balanced, as 42% of general surgery teaching faculty are women. Our finding that there were no differences in the "Overall value to learning" indicates that gender differences in faculty evaluations previously noted in the literature may not apply when there is a higher level of female representation.

Subtle gender biases within teaching evaluations are important to recognize because faculty evaluations by students and trainees are an important component in academic promotion. While the validity and fairness of teaching evaluations has been disputed,[2,9,10] they remain the most commonly used method of evaluating performance for faculty.[11] It is increasingly recognized that student evaluations of faculty affect promotion and tenure decisions [9] and are arguably the central

TABLE 2. Percentage of Male vs Female Faculty Scoring 6,7 by Category

				Univariate		Multivariate	
Category	**Subcategory**	**Male**	**Female**	**OR (95%CI)**	**p value**	**OR (95%CI)**	**p value**
Patient care	Deeply interested in patient care, making contributions to their management	1326 (84.7%)	1415 (89.6%)	0.64 (0.52-0.79)	<0.0001	1.5 (1.25-1.92)	<0.0001
	Sensitive to culture, age, gender, disability	1250 (81.2%)	1366 (87.4%)	0.62 (0.51-0.76)	<0.0001	1.62 (1.32-1.97)	<0.0001
	Compassionate to patient needs when performing procedures	1269 (84.8%)	1337 (88.5%)	0.72 (0.59-0.89)	0.003	1.39 (1.12-1.73)	0.003
Medical knowledge	Possesses excellent clinical acumen	1371 (86.9%)	1335 (84.8%)	1.2 (0.98-1.46)	0.08	0.82 (0.67-1.01)	0.06
	Displays investigatory and analytics thinking when discussing clinical problems	1371 (87.1%)	1348 (85.8%)	1.11 (0.91-1.37)	0.29	0.87 (0.71-1.07)	0.18
	Appropriately applies knowledge of his/her subject or specialty	1400 (88.7%)	1363 (86.6%)	1.21 (0.98-1.51)	0.07	0.80 (0.64-0.99)	0.04
Practice based learning and improvement	Demonstrates and fosters self-directed learning	1273 (82.8%)	1255 (81.5%)	1.09 (0.90-1.31)	0.38	0.90 (0.74-1.08)	0.25
	Ability to motivate, enthusiastic about subject	1321 (84.4%)	1309 (83.1%)	1.05 (0.87-1.27)	0.62	0.93 (0.76-1.12)	0.44
Interpersonal and communication skills	Presents material in organized, clear manner; summarizes major points and provides emphasis	1288 (82.9%)	1328 (85.1%)	0.85 (0.70-1.03)	0.10	1.17 (0.96-1.42)	0.13
	Devotes appropriate amounts of time and discussion in an open and friendly manner	1279 (82.0%)	1316 (83.9%)	0.89 (0.73-1.06)	0.17	1.11 (0.92-1.35)	0.26
	Is approachable and receptive to questions	1347 (85.9%)	1335 (84.7%)	1.10 (0.90-1.13)	0.37	0.88 (0.72-1.08)	0.23
	Effectively teaches and demonstrates technical, surgical and procedural skills	1245 (85.0%)	1233 (85.0%)	1.01 (0.82-1.23)	0.96	0.98 (0.80-1.21)	0.86
Professionalism	Enjoys teaching and is enthusiastic about subject	1349 (86.0%)	1342 (85.5%)	1.04 (0.65-1.27)	0.69	0.94 (0.77-1.15)	0.53
	Keeps appointments, punctual, does leave early, attentive during supervision	1299 (83.9%)	1390 (89.3%)	0.62 (0.51-0.77)	<0.0001	1.62 (1.31-2.00)	<0.0001
	Excellent role model as teacher and physician	1300 (83.1%)	1330 (84.4%)	0.91 (0.75-1.10)	0.33	1.07 (0.89-1.30)	0.47
	Commitment to the educational program	1346 (86.6%)	1357 (86.5%)	1.01 (0.82-1.24)	0.95	0.98 (0.80-1.21)	0.87
System based practice	Practices and teaches cost effective care	1200 (82.6%)	1173 (79.1%)	1.25 (1.04-1.50)	0.02	0.97 (0.80-1.17)	0.73
	Advocates for patients within the health care system	1250 (84.1%)	1267 (86.8%)	0.80 (0.65-0.99)	0.04	1.24 (1.01-1.53)	0.04
Overall	Overall value to your training	1306 (83.1%)	1261 (83.3%)	0.99 (0.82-1.19)	0.87	1.0 (0.83-1.21)	0.98

component to assessment of teaching.[12] As lack of academic progression is associated with increased burnout and abandonment of academic careers, any biases within student evaluations of faculty should be a focus of attention and intervention.[13] Older literature states women are less likely to be promoted than their male counterparts [8,9] and perceive gender as a limitation to career advancement in academic surgery.[10] Awareness and confrontation of these biases by student evaluators may help eliminate these limitations. Alternatively, some argue for diminishing the importance of these evaluations due to inherent biases over attempting to correct the biases themselves.[9] Our study suggests that differences in faculty evaluations by students dissipate with a higher female representation (even in a traditionally male dominated field like general surgery). As these biases may become less significant as equality is achieved, gender bias within evaluations may not limit the academic progression of male or female faculty.

Several investigations have shown that diversity in leadership is beneficial. In the business literature, gender diversity in leadership has a positive impact on performance.[14-16] In the academic world, overall learning and retention increases with the percentage of female faculty.[17] Traits of any good surgical teacher, regardless of gender, include intellectual attributes (clinical knowledge and clinical and technical competence), emotional intelligence (encouraging communication, being well-organized), and personality characteristics (being a positive role model, enthusiasm).[18,19] Given that faculty in our study had different strengths as teachers reflecting these traits, having a diverse faculty optimizes the learning environment and maximizes the likelihood of a great educational program.

There are a number of limitations to our study. This is a single institutional study, which limits the power of the findings. Our surgery program changed to MedHub 3 years ago, therefore only 2 full years of data was available. As the evaluations are anonymous, we are unable to evaluate the gender or postgraduate year of the evaluator. We were also unable to identify duplicates of the same pairing of faculty and resident evaluator. Additionally, we were unable to assess race, seniority, or professor rank. As our percentage of female faculty is twice the national average, it is difficult to determine how generalizable these findings are to all of general surgery. Future research will expand this investigation to multiple centers to determine if these findings persist in a more diverse group with variable leadership and amongst more or less gender-balanced programs.

CONCLUSION

In a general surgery residency program with a relatively gender-balanced faculty, there was no gender difference in the overall evaluation of faculty by residents. However, there were gender differences in specific domains. These findings suggest that gender balance in teaching faculty may help eliminate previously observed teaching evaluation bias in the traditionally male dominated fields. Future multi-institutional studies are needed to further investigate if similar findings persist amongst other general surgery departments of varying gender diversity.

REFERENCES

1. Kogan LR, Schoenfeld-Tacher R, Hellyer PW. Student evaluations of teaching: perceptions of faculty based on gender, position, and rank. *Teach Higher Educ*. 2010;15:623–636.
2. Hornstein HA. Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Educ*. 2017;4. https://doi.org/10.1080/2331186X.2017.1304016.
3. Miller J, Chamberlin A. Women are Teachers, men are Professors: a study of student perceptions. *Teach Sociol*. 2000;28:283–298.
4. Morgan HK, Purkiss JA, Porter AC, et al. Student evaluation of faculty physicians: gender differences in teaching evaluations. *J Womens Health*. 2016;25:453–456. https://doi.org/10.1089/jwh.2015.5475.
5. Fassiotto Magali, Li Jie, Maldonado Y. Female surgeons as counter Stereotype: the impact of gender perceptions on trainee evaluations of physician faculty. *J Surg Educ*. 2018;75:1140–1148. https://doi.org/10.1016/j.jsurg.2018.01.011.
6. Basow SA. Student evaluations of college professors: when gender matters. *J Educ Psychol*. 1995;87:656–665.
7. Bachen CM, McLoughlin MM, Garcia SS. Assessing the role of gender in college students' evaluations of faculty. *Commun Educ*. 1999;48:193–210. https://doi.org/10.1080/03634529909379169.
8. Feldmen K A. College students' views of male and female college teachers. Part II: evidence from students' evaluations of their classroom teachers. *Res Higher Educ*. 1993;34:151–191.
9. Haskell RE. Academic freedom, tenure, and student evaluation of faculty. *Educ Policy Anal Arch*. 1997;5:6. [S.l.] ISSN 1068−2341.
10. Cohen PA. Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies. *Rev Educ Res*. 1981;51:281–309.

11. Winkler JA. Faculty reappointment, tenure, and promotion: barriers for women. *Prof Geographer*. 2000;52:737–750.

12. Glick TH. How best to evaluate clinician–educators and teachers for Promotion? *Acad Med*. 2002; 77:392–397.

13. Lowenstein SR, Fernandez G, Crane LA. Medical school faculty discontent: prevalence and predictors of intent to leave academic careers. *BMC Med Educ*. 2007;7:37.

14. Haishan Li Peng Chen. Board gender diversity and firm performance: the moderating role of firm size. *Bus Ethics*. 2018;21.

15. Charles A, Dang R, Redor E. Board gender diversity and firm financial Performance: a quantile regression analysis. *Int Corp Governance Regul*. 2018: 15–55.

16. Lautenberger DM, Dandar VM, Raezer CL, Sloane RA. The state of women in academic medicine: the pipeline and pathways to leadership 2013–2014. *Assoc Am Med Colleges*. 2014: 1–17. https://store.aamc.org/downloadable/download/sample/sample_id/228/.

17. Sandler BR, Silverberg LA, Hall RM. The chilly classroom Climate: a guide to improve the education of women. Washington, DC: Natl Assn. Women Educ; 1996.

18. Sutkin G, Wagner E, Harris I, Schiffer R. What makes a good clinical teacher in Medicine? A review of the literature. *Acad Med*. 2008;83:452–466.

19. Sloan DA, Donnelly MB, Schwartz RW. The surgical clerkship: characteristics of the effective teacher. *Med Educ*. 1996;30:18–23.

SUPPLEMENTARY INFORMATION

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jsurg.2019.06.007.