



Published in final edited form as:

J Surg Educ. 2023 January ; 80(1): 102–109. doi:10.1016/j.jsurg.2022.08.018.

Gender bias in risk management reports involving physicians in training – a retrospective qualitative study

Elizabeth A. Andraska, MD MS^{*,a}, Amanda R. Phillips, MD MS^{*,a}, Sina Asaadi^{*}, Lisa Painter, DNP RN[§], Gregory Bump, MD^{||,‡}, Rabih Chaer, MD MS^{*,b}, Sara Myers, MD PhD^{†,b}

^{*} Division of Vascular Surgery, Department of Surgery, University of Pittsburgh, Pittsburgh, PA, USA.

[§] UPMC Corporate Risk Management, Pittsburgh, PA, USA.

^{||} UPMC Medical Education, University of Pittsburgh, Pittsburgh, PA, USA

[‡] Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA.

[†] Department of Surgery, University of Pittsburgh, Pittsburgh, PA, USA.

Abstract

Objective: Gender bias, which contributes to burnout and attrition of female medical trainees, may manifest as disparate workplace evaluations. Here, we explore gender-based differences in perceived competence and professionalism as described in an institutional electronic risk management reporting system.

Design: In this retrospective qualitative study, recurring themes were identified from anonymous entries reported to an electronic institutional risk management database from July 2014-July 2015 and July 2019-July 2020 using inductive methods. This electronic system is often used by hospital staff to document complaints against physicians under the pretext of poor patient care, regardless of whether an adverse event occurred. Two individuals independently coded entries. Themes were determined from event indicator codes (EIC) using Delphi methodology and compared between gender and specialty using bivariate statistics.

Setting: A multi-center integrated healthcare delivery system.

Participants: Risk management entries pertaining to physician trainees by hospital staff as written submissions to the institution's electronic risk management reporting system. Main outcomes included themes defined as: (1) lack of professionalism (i.e., delay in response, attitude, lack of communication), (2) perceived medical error, (3) breach of institutional protocol.

Corresponding Author Elizabeth Andraska, MD, MS, Division of Vascular Surgery, Department of Surgery, UPMC, University of Pittsburgh School of Medicine, 200 Lothrop St, F677 Presbyterian Hospital, Pittsburgh, PA 15213, Phone: 541-968-9627, andraskaea@upmc.edu.

^aDr. Andraska and Dr. Phillips contributed equally to writing, analysis, study design, and execution and would like to be considered as co-first authors.

^bDr. Myers and Dr. Chaer contributed equally to study design, execution, and dissemination and would like to be considered as co-senior authors.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Results: Of the 207 entries included for analysis, 52 entries identified men (25%) and 31 entries identified women (15%). The gender was not available in 124 entries and, therefore, categorized as ambiguous. The most common complaint about men involved a physician related EIC (n=12, 23%, EIC TX39) and the most common complaint about women involved a communication related EIC (n=7, 23%, EIC TX55). Eighty-eight (43%) entries involved medical trainees; 82 (40%) involved surgical trainees. Women were more often identified by their name only (n=8, 26% vs. n=3, 6%; $p<.001$). This finding was consistent in both medical (n=0, 0% vs. n=5, 31%; $p<.001$) and surgical (n=2, 7% vs. n=3, 25%; $p=.006$) specialties. In entries involving women, a lack of professionalism was most frequently cited (n=29, 94%). Entries identifying medical errors more frequently involved men (n=25, 48% vs. n=7, 23%; $p=.02$).

Conclusions: Gender-based differences exist in how hospital staff interpret trainees' actions and attitudes. These differences have consequences for training paradigms, perceptions of clinical competence, physician burnout and, ultimately, patient outcomes.

Keywords

Gender bias; trainees; burnout

Introduction

Existing studies focusing on the effect of gender disparities on attrition and burnout of women physicians-in-training have prompted initiatives from professional organizations such as the Association of American Medical Colleges (AAMC)¹ to mitigate inequalities that result from implicit and explicit gender bias.²⁻⁴ While trainees may experience or perceive gender bias in various ways, judgements regarding clinical competence seem to be a primary source. In nursing evaluations of residents' performance, female residents are more likely to receive critical feedback.^{4,5} These gender-based differences in perceptions of physician performance exist not only among healthcare providers, but also among patients; female physicians receive lower patient satisfaction scores than their male counterparts despite no differences in clinical outcomes.^{6,7}

Although previous investigations have highlighted gender discrepant performance evaluations,^{4,5,8} few have explored *how* feedback may be influenced by gender stereotypes. Specifically, expectations for men and women trainees may differ based on these stereotypes. Consequently, perceiving a trainee as 'competent,' may require different professional or behavioral parameters for women compared to men. In this qualitative study analyzing entries submitted anonymously by hospital staff to an electronic risk management portal designed to document incidents that threaten the integrity and efficiency of the healthcare system (e.g., medical errors, unprofessional behavior), we seek to explore gender-based differences in perceived competence and professionalism, hypothesizing that 1) female trainees will receive more complaints than male trainees, 2) these discrepancies will be more pronounced in surgical subspecialties where women may be underrepresented to a greater extent compared to medical specialties, and 3) there will be tangible differences in how women and men are perceived with respect to skill and professional communication.

Materials and Methods

Study design and data sources

In this retrospective qualitative study, entries submitted anonymously by hospital staff to an electronic event reporting system for a multi-center integrated health delivery system from July 1, 2014 through June 30, 2015 and July 1, 2019 through June 30, 2020 were examined for gender-based differences in the domains of 1) professionalism, 2) performance, and 3) breach of institutional protocol.⁹ Using rationale that over time, efforts have been made to reduce gender bias, entries from two discrete time periods were analyzed.

Reports were collected from an institutional online event reporting system. This electronic event reporting system is a tool that allows individuals to anonymously document details of medical errors and near-miss events for use in systems-based quality evaluations. This electronic system is also often used by hospital staff to document complaints against physicians under the pretext of poor patient care, regardless of whether an adverse event occurred. Once events are reported, they are inventoried at the institutional level according to an event indicator code (EIC) that classifies the entry for administrative purposes (see Supplement 1 for full list), and triaged to the risk management department.¹⁰ Delphi methodology¹¹ consensus among the investigators was established regarding EICs that would be included for analysis. Entries that were captured based on their EICs were then independently reviewed by two investigators to determine whether the entry clearly pertained to a physician trainee (i.e., resident or fellow). The gender distribution of subjects of the entries was compared to the proportion of self-identified male and female trainees in each of the respective years (obtained from the institutional UPMC Graduate Medical Education (GME) office, and nationally, from the national GME website¹²) in order to control for selection bias. This study adheres to the Standards for Reporting in Qualitative Research (SRQR) guidelines¹³ and was approved by the UPMC quality review committee (QRC ID 2777) with a waiver of informed consent due to minimal risk to patients.

Data, coding, and thematic analysis

Subject training status, specialty (medical vs. surgical), and appellation (e.g., professional title vs. name) were abstracted for each entry. For all entries, efforts were made to identify the resident or fellow gender. Unfortunately, many entries were truly ambiguous, often including statements such as ‘the resident did the following...’. In such circumstances, no details were present to help clarify the residents’ gender, and the subject’s gender was labelled ‘ambiguous’. A preliminary coding scheme was created by analyzing a subset of the entries using *a priori* codes for themes based off of existing studies investigating interprofessional workplace conflicts experienced by women surgeons.¹⁴ These themes were then revised on the basis of clarity and relevance until consensus had been established.¹⁵ The entries were examined independently by two coders for the following themes which have been modified from previously reported domains: (1) lack of professionalism (i.e., delay in response, attitude, lack of communication), (2) perceived medical error, (3) breach of institutional protocol.¹⁴

Cohen's Kappa was used to calculate interrater reliability for each coded item in the dataset and yielded an average of 68% agreement. Bivariate statistics (Mann-Whitney U test) using STATA 16SE statistical software (StataCorp; College Station, TX) were used to compare gender-based differences in year of submission and trainees' subspecialty (medicine vs. surgery). A P value $<.05$ was considered the threshold for statistical significance. As entries necessarily represent only the experiences that individuals chose to report rather than all that may have occurred, descriptive statistics were used for thematic analysis.

Results

Influence of gender on risk management complaints

7,223 entries were submitted to the risk management portal during the study periods. 2,013 entries were included with relevant EIC. 235 of these entries identified a trainee as the subject. An additional 28 entries were excluded due to an incomplete event form, resulting in 207 entries included for analysis (Figure 1). Among reported entries, there were no differences in trainee gender (2014–2015 women = 17, 55% vs. 2019–2020 women = 14, 45%; $p=.12$) or trainee specialty (2014–2015 surgical trainees = 42, 51% vs. 2019–2020 surgical trainees = 40, 49%; $p=.37$) between time periods. Of the 83 (40%) entries that identified the trainee's gender, 31 (37%) identified women while 52 (63%) identified men. Local GME records for the UPMC healthcare system reveal that during the specified study periods, 46% trainees were women and 54% men (none self-reported as non-binary). 82 (40%) of the entries concerned trainees from surgical specialties with the remainder involving medical subspecialty trainees. During the study periods, 24% of trainees at UPMC were completing surgical subspecialty training. Among the entries concerning surgical subspecialty trainees, 12 (39%) were identified women and 28 (54%) identified men. The gender-composition of institutional GME trainees mirrored the gender distribution within entries (women medical trainees: 52% vs. 49%, women surgical trainees: 39% vs. 37%). Details regarding subject gender and specialty are displayed in Table 1.

Appellations

Men and women were referred to by their professional title with similar frequency (men=42, 81% vs. women=19, 61% vs. ambiguous=78, 63%; $p=.05$). Women trainees, however, were more likely to be referred to by name only (men=3, 6% vs. women=8, 26% vs. ambiguous=0, 0%; $p<.001$). This finding applied to trainees in both medical (men=0, 0% vs. women=5, 31% vs. ambiguous=0, 0%; $p<.001$) and surgical (men=2, 7% vs. women=3, 25% vs. ambiguous=0, 0%; $p=.006$) specialties. When comparing women in a medical specialty to women in a surgical specialty, there were no differences in appellation (Table 2).

Gender-based differences in complaint type

The remaining analysis compares men versus women directly. The majority of all submissions described a lack of professionalism (men=42, 81% vs. women=29, 94%; $p=.11$). 32 (39%) entries described trainees committing medical error, with more men than women being cited (men=25, 48% vs. women=7, 23%; $p.02$). Among the 19 (23%) of entries that described a breach in institutional protocol (men=12, 23% vs. women=7, 23%; $p=.96$). Among complaints that described a lack of professionalism, men and women

had similar rates of entries describing delay in care (men=23, 44% vs. women=16, 52%; $p=.52$) and unacceptable attitude (men=21, 40% vs. women=10, 32%; $p=.46$). On the other hand, women were more frequently cited for a lack of communication (men=11, 21% vs. women=17, 55%, $p=.002$). While more surgical trainees were cited for lack of professionalism, in general (medical=27, 75% vs. surgical=37, 93%; $p=.04$), there was no difference in the frequency of these complaints when comparing women surgical trainees to women medical trainees (Table 3). When comparing entries where the gender was ambiguous, most themes landed in between the reported proportions for men and women (Supplement 2). There were two notable deviations from this trend. A lack of professionalism citing attitude specifically, was lower in ambiguous entries compared to both male and female entries. Additionally, lack of communication was much lower in male entries compared to both women and ambiguous entries. Men are less likely to be professionally cited for communication discrepancies than their counterparts.

19 (23%) entries used an adjective to describe the trainee or trainee's behavior. Examples of ways that were used to describe men included "unprofessional", "mean", and "insulting". Examples of adjectives that were used to describe women included "insensitive", "arrogant", "demeaning", and "set in her ways" (Table 4).

Discussion

This retrospective qualitative analysis of entries submitted to an electronic institutional risk management portal to document medical errors and near-miss events demonstrated important gender-based differences in entry type and content. While overall, the proportion of entries about male and female trainees mirrored the gender composition of GME nationally and institutionally, women were more frequently criticized for lack of communication whereas complaints about male trainees were more often related to perceived medical errors. Female trainees were more likely than their male counterparts to be referred to by their name alone.

This study expands on the current literature regarding gender bias in medical training and reflects how men and women may differ with respect to the type of criticism they receive.^{8,16} We found that women were more likely than men to be criticized for violating their perceived station. For example, women were cited as "inappropriate", "condescending", "demeaning", and "set in her ways" whereas men were described as being "abrupt", "mean", and "insulting". Societal hierarchy places women in a submissive and inferior role compared to men. In an environment such as the hospital, where hierarchy is reinforced, women physicians in training are constantly violating their expected station. Critiques such as being "condescending" or "demeaning" suggest that there is an expectation that women trainees are to be lower in the hierarchy. This finding is consistent with other investigations that show counter-normative behavior can lead to unequal penalties.^{4,5,17} For example, nurses tend to evaluate female trainees lower than male trainees with regards to both ability and work ethic even though there are no differences in competence measured by in-service exams.⁴

Our data indicating that women who deviate from stereotypical behaviors are often considered unprofessional is corroborated by existing data. Fassio *et al* found that in their evaluations, women physicians scored lower independent of performance. It is theorized that this phenomenon is especially pronounced in non-primary care fields, such as surgery and its subspecialties, where women providers represent a smaller percentage of the workforce.^{18,19} While not statistically significant, surgical trainees in the current study were slightly overrepresented in the reporting system relative to medical trainees. This may be due to a general assumption that surgeons exhibit more masculine traits and thus female surgical trainees violate implicit biases more than medical trainees. Thus, they may be more vulnerable to critiques on professionalism in this setting. Regarding women as unprofessional on the basis of non-adherence to gender stereotypes^{18,20,21} threatens the quality of our medical training²² and promotes microaggressions that contribute to physician burnout.² Furthermore, dysfunctional medical teams have been associated with poor patient outcomes and higher rates of medical errors.^{23–25} Additionally, one could speculate that it encourages the distorted gender-specific perception of physicians and surgeons that may underwrite appraisals of clinical competence, or lack thereof, salary discrepancies, inequitable promotion criteria, etc.

Female physicians are often tasked with balancing a perceived need to be sensitive and caring with the occupational need to be autonomous and assertive.²¹ In fact, it has been reported that patients value traditionally feminine traits such as empathy, shared decision making, and being approachable.²⁶ This results in discordant expectations from hospital staff and inconsistencies in what is deemed to be professional behavior for females, which is reflected in our study in the adjectives used to describe female trainees. Mueller *et al* described this finding in physician evaluations of trainees, insofar as female trainees received conflicting feedback with respect to professionalism.⁸ Our study found that while women were more frequently cited for unprofessional attitudes or behaviors, this did not differ based on their subspecialty (i.e., medicine vs. surgery). Less than a quarter of the complaints against female trainees, however, involved tangible medical error, as also shown in our study, indicating gender discordance between competence and perceived professionalism. The latter may provide insight into structural bias that may contribute to findings that women are less likely to meet milestones, attain faculty positions, and be retained in the field.^{17,27–29} Our findings suggest asymmetric expectations of men and women trainees by hospital staff, which may be influenced by ubiquitous gender-based stereotypes and not reflective of clinical competence. In fact, in our study, trainees who were men were more likely to be reported for medical errors compared to women. This is consistent with emerging literature to suggest improved patient outcomes when under the care of a female physician.³² In the present study, data was obtained from a reporting system aimed to capture medical errors. Despite this, women were consistently more likely to be reported for professionalism, whereas men more often were reported for medical error. More studies are necessary to evaluate the influence of gender on provided care and the subsequent consequences to the patient.

In our study, women were more likely to be referred to by name only, in contrast to their professional title. There is a significant amount of literature describing bias when introducing speakers at international meetings.^{30,31} Duma *et al* evaluated over 2,500 videos

of presentations at international oncology meetings and reported a significant bias relating to gender disparities in speaker introductions. Women were far less likely to be introduced by their professional title compared to men. There is a consistent bias in medicine to not acknowledge the professional status of women. In this study, we found this on the level of hospital staff trainee interactions. Refusing to properly address one's professional station can potentially further perpetuate gender bias in the field.

Several limitations exist for this retrospective qualitative study. The majority of entries submitted lacked data regarding trainee specialty or gender, which may have biased our results which reflected only reports that specifically stated these data. Entries to the risk management reporting system were used as a surrogate for attitudes towards trainees and may not be completely representative of one's attitude. The study also was entirely anonymous, limiting the authors' ability to discern details of the individuals who initiated the complaints. While this study is limited by the anonymity of the entry writers, it will be important to more deeply evaluate the roles of other health leaders in the future. Charge nurses, nurse managers, health directors, all represent leaders in health care who may not be viewed as aligning with social gender expectations. Further, this study suggests that there are systemic misperceptions based on gender that are perpetuated in health care and may result from misalignment with gender expectations and role in medical hierarchy. This is borne in microaggressions and gender specific discrimination and warrants further study to evaluate its implications on provider mental health, burnout, quality of care provided, and medical systems as a whole. Finally, the study did not include trainees who identify as non-binary, which excludes an important training experience. We also were unable to obtain race, ethnicity, or sexual orientation details on the cited trainees, so the interpretations outlined here do not account for intersectionality of minority experience. Despite these limitations, the findings from this study demonstrate the ongoing trend in gender bias in medical education.

Conclusion

Gender-based differences exist in how hospital staff interpret trainees' actions and attitudes. These differences have consequences for training paradigms, perceptions of clinical competence, physician burnout, and, ultimately, patient outcomes. Future studies are needed to evaluate and eliminate the effect of workplace gender bias on trainee clinical and professional development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Source of support:

This research was supported in part by grant T32HL98036 from the National Heart, Lung, and Blood Institute (Andraska, Phillips). The University of Pittsburgh holds a Physician-Scientist Institutional Award from the Burroughs Wellcome Fund (Andraska).

References

1. Redford GPB. AAMC launches new initiative to address and eliminate gender inequities. AAMC.
2. Dahlke AR, Johnson JK, Greenberg CC, et al. Gender Differences in Utilization of Duty-hour Regulations, Aspects of Burnout, and Psychological Well-being Among General Surgery Residents in the United States. *Ann Surg*. 2018;268(2):204–211. doi:10.1097/SLA.0000000000002700 [PubMed: 29462009]
3. Gjerberg E, Kjølørød L. The doctor-nurse relationship: how easy is it to be a female doctor cooperating with a female nurse? *Soc Sci Med*. 2001;52(2):189–202. doi:10.1016/s0277-9536(00)00219-7 [PubMed: 11144775]
4. Brucker K, Whitaker N, Morgan ZS, et al. Exploring Gender Bias in Nursing Evaluations of Emergency Medicine Residents. *Acad Emerg Med Off J Soc Acad Emerg Med*. 2019;26(11):1266–1272. doi:10.1111/acem.13843
5. Galvin SL, Parlier AB, Martino E, Scott KR, Buys E. Gender Bias in Nurse Evaluations of Residents in Obstetrics and Gynecology. *Obstet Gynecol*. 2015;126 Suppl:7S–12S. doi:10.1097/AOG.0000000000001044 [PubMed: 26375558]
6. Rogo-Gupta LJ, Haunschild C, Altamirano J, Maldonado YA, Fassiotto M. Physician Gender Is Associated with Press Ganey Patient Satisfaction Scores in Outpatient Gynecology. *Women's Heal issues Off Publ Jacobs Inst Women's Heal*. 2018;28(3):281–285. doi:10.1016/j.whi.2018.01.001
7. Hall JA, Blanch-Hartigan D, Roter DL. Patients' satisfaction with male versus female physicians: a meta-analysis. *Med Care*. 2011;49(7):611–617. doi:10.1097/MLR.0b013e318213c03f [PubMed: 21478782]
8. Mueller AS, Jenkins TM, Osborne M, Dayal A, O'Connor DM, Arora VM. Gender Differences in Attending Physicians' Feedback to Residents: A Qualitative Analysis. *J Grad Med Educ*. 2017;9(5):577–585. doi:10.4300/JGME-D-17-00126.1 [PubMed: 29075375]
9. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol*. 2014;67(4):401–409. doi:10.1016/j.jclinepi.2013.12.002 [PubMed: 24581294]
10. Painter LM, Kidwell KM, Kidwell RP, et al. Do Written Disclosures of Serious Events Increase Risk of Malpractice Claims? One Health Care System's Experience. *J Patient Saf*. 2018;14(2):87–94. doi:10.1097/PTS.0000000000000178 [PubMed: 25831069]
11. Beiderbeck D, Frevel N, von der Gracht HA, Schmidt SL, Schweitzer VM. Preparing, conducting, and analyzing Delphi surveys: Cross-disciplinary practices, new directions, and advancements. *MethodsX*. 2021;8:101401. doi:10.1016/j.mex.2021.101401 [PubMed: 34430297]
12. ACGME. Number of Active Residents, by Type of Medical School, GME Specialty, and Sex.
13. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med*. 2014;89(9):1245–1251. doi:10.1097/ACM.0000000000000388 [PubMed: 24979285]
14. Dossett LA, Vitous CA, Lindquist K, Jagsi R, Telem DA. Women Surgeons' Experiences of Interprofessional Workplace Conflict. *JAMA Netw open*. 2020;3(10):e2019843. doi:10.1001/jamanetworkopen.2020.19843 [PubMed: 33030551]
15. Stuckey HL. The second step in data analysis: Coding qualitative research data. *J Soc Heal Diabetes*. 2015;3(01):7–10.
16. Dayal A, O'Connor DM, Qadri U, Arora VM. Comparison of Male vs Female Resident Milestone Evaluations by Faculty During Emergency Medicine Residency Training. *JAMA Intern Med*. 2017;177(5):651–657. doi:10.1001/jamainternmed.2016.9616 [PubMed: 28264090]
17. Bear JB, Cushenbery L, London M, Sherman GD. Performance feedback, power retention, and the gender gap in leadership. *Leadersh Q*. 2017;28(6):721–740.
18. Fassiotto M, Li J, Maldonado Y, Kothary N. Female Surgeons as Counter Stereotype: The Impact of Gender Perceptions on Trainee Evaluations of Physician Faculty. *J Surg Educ*. 2018;75(5):1140–1148. doi:10.1016/j.jsurg.2018.01.011 [PubMed: 29402668]
19. Olumolade OO, Rollins PD, Daignault-Newton S, George BC, Kraft KH. Closing the Gap: Evaluation of Gender Disparities in Urology Resident Operative Autonomy and Performance. *J Surg Educ*. November 2021. doi:10.1016/j.jsurg.2021.10.010

20. Barnes KL, McGuire L, Dunivan G, Sussman AL, McKee R. Gender Bias Experiences of Female Surgical Trainees. *J Surg Educ.* 2019;76(6):e1–e14. doi:10.1016/j.jsurg.2019.07.024 [PubMed: 31601487]
21. Chen Y-W, Chang DC, Armstrong KA, Kelleher CM. “Women as Deficit”: Re-evaluating Interventions to Establish Gender Equity. *Ann Surg.* 2022;275(1):37–38. doi:10.1097/SLA.0000000000005170 [PubMed: 34387210]
22. Myers SP, Hill KA, Nicholson KJ, et al. A qualitative study of gender differences in the experiences of general surgery trainees. *J Surg Res.* 2018;228:127–134. doi:10.1016/j.jss.2018.02.043 [PubMed: 29907201]
23. Herzberg S, Hansen M, Schoonover A, et al. Association between measured teamwork and medical errors: an observational study of prehospital care in the USA. *BMJ Open.* 2019;9(10):e025314. doi:10.1136/bmjopen-2018-025314
24. Risser DT, Rice MM, Salisbury ML, Simon R, Jay GD, Berns SD. The potential for improved teamwork to reduce medical errors in the emergency department. The MedTeams Research Consortium. *Ann Emerg Med.* 1999;34(3):373–383. doi:10.1016/s0196-0644(99)70134-4 [PubMed: 10459096]
25. Horwitz SK, Horwitz IB, Barshes NR. Addressing dysfunctional relations among healthcare teams: improving team cooperation through applied organizational theories. *Adv Health Care Manag.* 2011;10:173–197. doi:10.1108/s1474-8231(2011)0000010017 [PubMed: 21887945]
26. Chen H, Pierson E, Schmer-Galunder S, et al. Gender Differences in Patient Perceptions of Physicians’ Communal Traits and the Impact on Physician Evaluations. *J Womens Health (Larchmt).* 2021;30(4):551–556. doi:10.1089/jwh.2019.8233 [PubMed: 32857642]
27. Mueller CM, Gaudilliere DK, Kin C, Menorca R, Girod S. Gender disparities in scholarly productivity of US academic surgeons. *J Surg Res.* 2016;203(1):28–33. doi:10.1016/j.jss.2016.03.060 [PubMed: 27338531]
28. Khoushhal Z, Hussain MA, Greco E, et al. Prevalence and Causes of Attrition Among Surgical Residents: A Systematic Review and Meta-analysis. *JAMA Surg.* 2017;152(3):265–272. doi:10.1001/jamasurg.2016.4086 [PubMed: 27973673]
29. Lyons NB, Bernardi K, Huang L, et al. Gender Disparity in Surgery: An Evaluation of Surgical Societies. *Surg Infect (Larchmt).* 2019;20(5):406–410. doi:10.1089/sur.2018.220 [PubMed: 30892131]
30. Huang CC, Lapen K, Shah K, et al. Evaluating Bias in Speaker Introductions at the American Society for Radiation Oncology Annual Meeting. *Int J Radiat Oncol Biol Phys.* 2021;110(2):303–311. doi:10.1016/j.ijrobp.2020.12.027 [PubMed: 33373658]
31. Files JA, Mayer AP, Ko MG, et al. Speaker Introductions at Internal Medicine Grand Rounds: Forms of Address Reveal Gender Bias. *J Womens Health (Larchmt).* 2017;26(5):413–419. doi:10.1089/jwh.2016.6044 [PubMed: 28437214]
32. Tsugawa Y, Jena AB, Figueroa JF, Orav EJ, Blumenthal DM, Jha AK. Comparison of Hospital Mortality and Readmission Rates for Medicare Patients Treated by Male vs Female Physicians. *JAMA Intern Med.* 2017;177(2):206–213. doi:10.1001/jamainternmed.2016.7875 [PubMed: 27992617]

Highlights

- Among hospital staff related complaints about graduate medical trainees, women are more often cited for lack of communication while men are more often cited for medical error
- Women trainees are more often referred to by their name whereas trainees who are men are referred to by their professional title
- Adjectives to describe women in written complaints often center around a violation of perceived social hierarchy

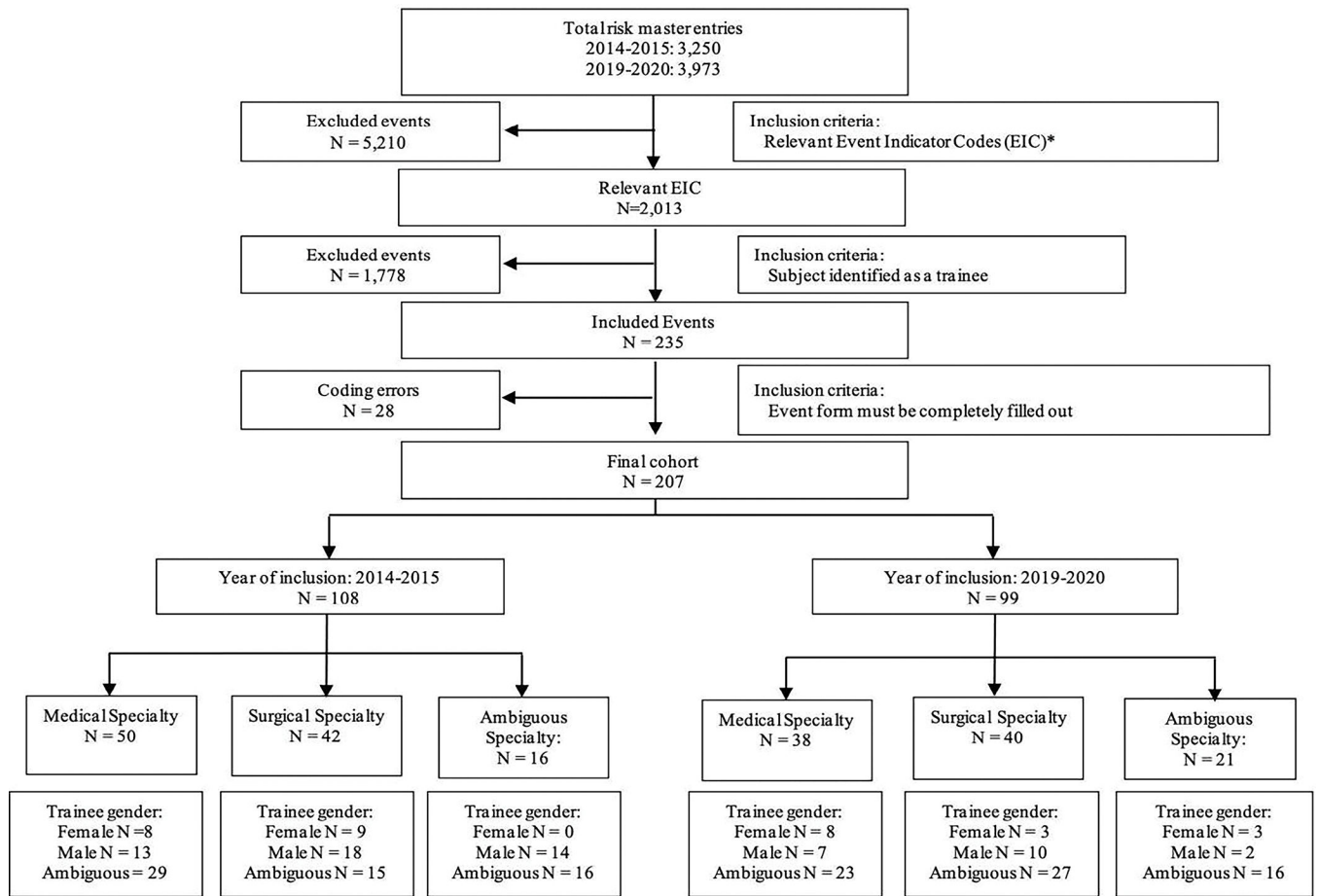


Figure 1.
Flow chart depicting the sample cohort.
^aEvent indicator codes included in Appendix A

Table 1.

Gender and specialty differences in entries and corresponding institutional and national data during the study periods.

Specialty	Men (n=52)	Women (n=31)	Ambiguous (n=124)	Total (N=207)	<i>p</i> value	UPMC women	ACGME women
Medical	20 (39)	16 (52)	52 (42)	88 (42.5)	.49	49.0%	48.2%
Surgical	28 (54)	12 (39)	42 (34)	82 (39.6)	.05	35.9%	31.4%
Ambiguous	4 (8)	3 (10)	30 (24)	37 (17.9)	.02		

Table 2.

Difference in titles used in reference to trainees in written entries by gender and specialty.

All entries	Men (n=52)	Women (n=31)	Ambiguous (n=124)	Total (N=207)	p value
Professional title	42 (81)	19 (61)	78 (63)	139 (67)	.05
Name only	3 (6)	8 (26)	0 (0)	11 (5)	<.001
Medical specialty	Men (n=20)	Women (n=16)	Ambiguous (n=52)	Total (n=88)	p value
Professional title	18 (90)	10 (63)	39 (75)	67 (76)	.15
Name only	0 (0)	5 (31)	0 (0)	5 (6)	<.001
Surgical specialty	Men (n=28)	Women (n=12)	Ambiguous (n=42)	Total (n=82)	p value
Professional title	21 (75)	8 (67)	17 (41)	46 (56)	.01
Name only	2 (7)	3 (25)	0 (0)	5 (6)	.006
Women only	Medical specialty (N=16)	Surgical specialty (N=12)	Ambiguous specialty (N=3)	Total (n=31)	p value
Professional title	10 (63)	8 (67)	1 (33)	19 (61)	.56
Name only	5 (31)	3 (25)	0 (0)	8 (26)	.52

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Difference in themes of entry by gender and specialty^a

Theme, by gender^b	Men (N=52)	Women (N=31)	Total (N=83)	P value
Lack of professionalism				
Any	42 (81)	29 (94)	71 (86)	.11
Delay	23 (44)	16 (52)	39 (47)	.52
Attitude	21 (40)	10 (32)	31 (37)	.46
Lack of communication	11 (21)	17 (55)	28 (34)	.002
Perceived medical error	25 (48)	7 (23)	32 (39)	.02
Breach in institutional protocol	12 (23)	7 (23)	19 (23)	.96
Theme, by specialty^c	Medical specialty (N=36)	specialtySurgical specialty (N=40)	Total (N=76)	P value
Lack of professionalism				
Any	27 (75)	37 (93)	64 (84)	.04
Delay	14 (39)	25 (63)	39 (51)	.04
Attitude	13 (36)	12 (30)	25 (33)	.57
Lack of communication	15 (42)	10 (25)	25 (33)	.12
Perceived medical error	17 (47)	13 (33)	30 (40)	.19
Breach in institutional protocol	4 (11)	15 (38)	19 (25)	.008
Theme, women only^d	Medical specialty (N=16)	Surgical specialty (N=12)	Total (N=28)	P value
Lack of professionalism				
Any	14 (88)	12 (100)	26 (93)	.20
Delay	8 (50)	8 (67)	16 (57)	.28
Attitude	6 (38)	2 (17)	8 (29)	.23
Lack of communication	10 (63)	5 (42)	15 (54)	.27
Perceived medical error	5 (31)	2 (17)	7 (25)	.38
Breach in institutional protocol	2 (13)	5 (42)	7 (25)	.08

^aThemes are not mutually exclusive

^bEntries with no identifiable gender excluded (n=124)

^cEntries with no identifiable specialty excluded (n=7)

^dEntries with identified gender and specialty only

Table 4.

Differences in descriptive words used to describe trainees in entries by gender.

	Men (n=52)	Women (n=31)	Total (n=83)	<i>p</i> value
Use of a descriptive word or phrase	12 (23)	7 (23)	19 (23)	.83
Examples	Lack of empathy, extremely rude, abrupt, unprofessional, harassed, not attentive, dismissive, mean, violated, blew her off, passing the buck, mistreatment, insulting, humiliating	Impatient, inappropriate, insensitive, arrogant, handled roughly, condescending, complaining, demeaning, arrogant, set in her ways, abrupt		

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript