



Does Applicant Gender Have an Effect on Standardized Letters of Evaluation Obtained During Medical Student Emergency Medicine Rotations?

Jessica Andrusaitis, MD, MS¹, Clelia Clark, MBA¹, Soheil Saadat, MD, MPH, PhD², John Billimek, PhD¹, Sara Paradise, MD², Alisa Wray, MD, MAEd², Warren Wiechmann, MD, MBA², Shannon Toohey, MD, MAEd² , and Megan Boysen-Osborn, MD, MHPE^{1,2} 

ABSTRACT

Background: The standardized letter of evaluation (SLOE) in emergency medicine (EM) is one of the most important items in a student's application to EM residency and replaces narrative letters of recommendation. The SLOE ranks students into quantile categories in comparison to their peers for overall performance during an EM clerkship and for their expected rank list position. Gender differences exist in several assessment methods in undergraduate and graduate medical education. No authors have recently studied whether there are differences in the global assessment of men and women on the SLOE.

Objectives: The objective of this study was to determine if there is an effect of student gender on the outcome of a SLOE.

Methods: This was a retrospective observational study examining SLOEs from applications to a large urban, academic EM residency program from 2015 to 2016. Composite scores (CSs), comparative rank scores (CRSs), and rank list position scores (RLPSs) on the SLOE were compared for female and male applicants using Mann-Whitney U-test.

Results: From a total 1,408 applications, 1,038 applicants met inclusion criteria (74%). We analyzed 2,092 SLOEs from these applications. Female applicants were found to have slightly lower and thus better CRSs, RLPSs, and CSs than men. The mean CRS for women was 2.27 and 2.45 for men ($p < 0.001$); RLPS for women was 2.32 and 2.52 for men ($p < 0.001$) and CS was 4.59 for women and 4.97 for men ($p < 0.001$).

Conclusions: Female applicants have somewhat better performance on the EM SLOE than their male counterparts.

From the ¹University of California Irvine, CA; and ²Department of Emergency Medicine, University of California, Irvine, Orange, CA. Received June 26, 2019; revision received August 21, 2019; accepted August 28, 2019.

Presented at American College of Emergency Physicians Scientific Assembly, San Diego, CA, October 1, 2018.

The authors have no relevant financial information or potential conflicts of interest to disclose.

Author contributions: MBO and JA conceived the study; JA and MBO designed the experiment and obtained IRB approval; MBO supervised the conduct of the trial and data collection; JA and CC performed the data collection; JB and SS provided statistical advice on study design; and JB, JA, SS, ST, SP, and MBO analyzed the data; JA, CC, MBO, SP, ST, AW, WW, and CC drafted the manuscript; and all authors contributed substantially to its revision. MBO takes responsibility for the paper as a whole.

Dr. Andrusaitis is with the Kaiser Permanente Emergency Medicine Residency, San Diego, CA. Dr. Clark is currently with the University of California at Davis, Sacramento, CA.

Supervising Editor: Daniel P. Runde, MD, MME.

Address for correspondence and reprints: Megan Boysen-Osborn, MD, MHPE, e-mail: mbo@uci.edu.

AEM EDUCATION AND TRAINING 2020;4:18–23.

In 1995, the Council of Residency Directors in Emergency Medicine (CORD-EM) established a task force to create a standardized letter of recommendation (SLOR).¹ Renamed the standardized letter of evaluation (SLOE) in 2013², the intention of the SLOR/SLOE was to make a standardized evaluation that could stratify students into quantiles based on performance during an emergency medicine (EM) rotation.³

The SLOE replaces traditional narrative letters of recommendation for applicants to EM residency programs. Program directors (PDs) regard the SLOE as the most important item used for selecting applicants to interview.^{4–6} Students typically include two or more SLOEs in their electronic residency service application (ERAS) to EM residency programs.⁷ While there are many benefits of the SLOE,^{1,3,9,10} it is unclear whether the SLOE can predict future performance.^{11–13}

The SLOE ranks students in various aspects of their EM clerkship performance (e.g., work ethic, team work, anticipated level of guidance needed).⁹ The global assessment section (section C) of the SLOE places the student into a quantile for two questions. Question C1 asks: “Compared to other EM residency candidates you have recommended in the last academic year the candidate is: top 10%, top third, middle third, lower third.” Question C2b asks: “How highly would you estimate the candidate will reside on your rank list: top 10%, top third, middle third, lower third, unlikely to be on our rank list.”

The SLOE global assessment scores may be inflated, with most students clustering in the higher categories.^{3,4,14–17} A study from 2004 found that female authors assessing female students tended to give higher global assessment scores on the SLOR than female authors assessing males or male authors assessing males or females.¹⁸ Currently, most SLOEs are written by a group of authors, including the clerkship and residency leadership, so the influence of author gender may be less relevant;^{9,15} however, a difference in SLOE scores for female and male students may still exist. Studies from undergraduate and graduate medical education have demonstrated gender differences associated with several assessment methods.^{18–28} For example, men in EM residencies achieve higher milestone assessments than women at graduation.²⁰

With the revised format of the SLOE² and the tendency for more SLOEs to be completed by groups of EM residency and clerkship leaders,^{9,15} rather than individual faculty, we sought to determine if there is any difference in the SLOE global assessment of

female versus male applicants. A recent study by Li and colleagues³⁰ did not find any gender bias associated with the words used to evaluate female applicants; however, Li et al. did not study quantitative findings from SLOE rankings. Given the significance of the SLOE for applicants to EM residency programs, it is important to evaluate its objectivity, observing if any implicit bias exists among SLOE writers. As the primary outcome for our study, we sought to determine whether there was a difference in performance on SLOE global assessment scores between self-identified female and male applicants.

METHODS

We performed a retrospective review of SLOEs from applicants to the University of California at Irvine (UCI) EM residency program through the ERAS during the 2014–2015 and 2015–2016 application cycles. The UCI EM residency program is 3-year program in an urban academic medical center in the greater Los Angeles metropolitan area.

We included records for applicants from all Liaison Committee for Medical Education Doctor of Medicine (MD) granting schools, excluding Puerto Rico and Canada. Two trained, nonblinded abstractors collected data using a standardized data abstraction form. The abstractors recorded Association of American Medical Colleges (AAMC) identification number, applicant gender, application year, medical school attended, United States Medical Licensing Examination (USMLE) step 1 score, step 2 clinical knowledge (CK) score, and data from each SLOE. From the SLOE, we collected the following: rotation location, date of rotation, author type (e.g., PD, clerkship director [CD], group, or faculty), number of letters authored in the previous year, comparative rank score (CRS), and rank list position score (RLPS). We did not record the gender of the letter writer, as most SLOEs were group SLOEs. The CRS and RLPS are the quantile answers (e.g. “top third”) to questions C1 and C2b, respectively, which are described in the introduction. The data abstractors and senior investigator held periodic meetings to discuss data abstraction and answer questions, which were resolved via consensus. The senior investigator sampled 5% of records to ensure quality. We calculated Cohen’s kappa for agreement. The senior investigator reviewed all final data for omissions and accuracy prior to analysis. The data were stored in a secure database, REDCap.

We screened and recorded data from SLOEs from all 4-week or 1-month traditional EM rotations (i.e., not pediatric EM or EM ultrasound). For the study analysis, we excluded subjects if they did not have any SLOEs in their applicant file. We excluded SLOEs if they were not written by an author who had written more than five SLOEs in the previous year, unless they predicted that they would write more than five in the current academic year. We excluded SLOEs that were not written by a faculty group, PD, CD, or any combination of these. We excluded SLOEs if they did not have complete data available.

For the CRS, we assigned a value to the following categories: top 10% (1 point), top third (2 points), middle third (3 points), and lower third (4 points). For the RLPS, we assigned the same values to each category and assigned 5 points to the “unlikely to be on our rank list” category. We calculated a composite score (CS) for each SLOE, which is the sum of the CRS and RLPS. The best CS possible is 2, which corresponds to “top 10%, top 10%,” whereas the worst possible CS is 9, which corresponds to “lower third, unlikely to be on our rank list.” We wanted to detect a difference of 15% in CSs for female and male applicants, which corresponds to an approximate single quantile difference (top third versus middle third) for a middle third, middle third candidate. Assuming a reliability of 50%, the required sample size was 167 participants in each group (334 total participants) for 80% power and significance of $p = 0.05$.

We describe the distribution of scores for the CRS and RPLS for female versus male applicants. We compared the mean CRS, RLPS, and CS for all SLOEs from students who identified as female versus those who identified as male, using the Mann-Whitney U-test. We used IBM SPSS Statistics for Windows, version 25.0, to analyze data. We obtained institutional review board (IRB) approval from UCI before the study commenced. The IRB did not require subject consent.

RESULTS

There were 4,066 total EM applicants from U.S. MD-granting schools in 2015 and 2016.³¹ The UCI Emergency Medicine Residency Program received 640 applications in 2015 and 768 applications in 2016, totaling 1,408 applicants. Of these, 1,053 were applicants from U.S. MD-granting schools. For the study analysis, we excluded 15 applicants because they did not have any

SLOEs in their file that met inclusion criteria. We included 1,038 applications (74% of total applications) in the study analysis. All students reported either male or female gender. Most students included in the study identified as men ($n = 685$, 66%), which was similar to the percentage of male residents in EM in 2015 (63.1%).³¹ Agreement between the senior investigator and data abstractors was substantial ($\kappa = 0.919$ –1.000).

Applicants had one ($n = 132$, 13%), two ($n = 644$, 62%), or three ($n = 262$, 25%) SLOEs in their files, totaling 2,206 SLOEs. We excluded 114 of these SLOEs because they had incomplete data; were not written by a group, PD, CD, or combination; or were not written by an author who had written more than

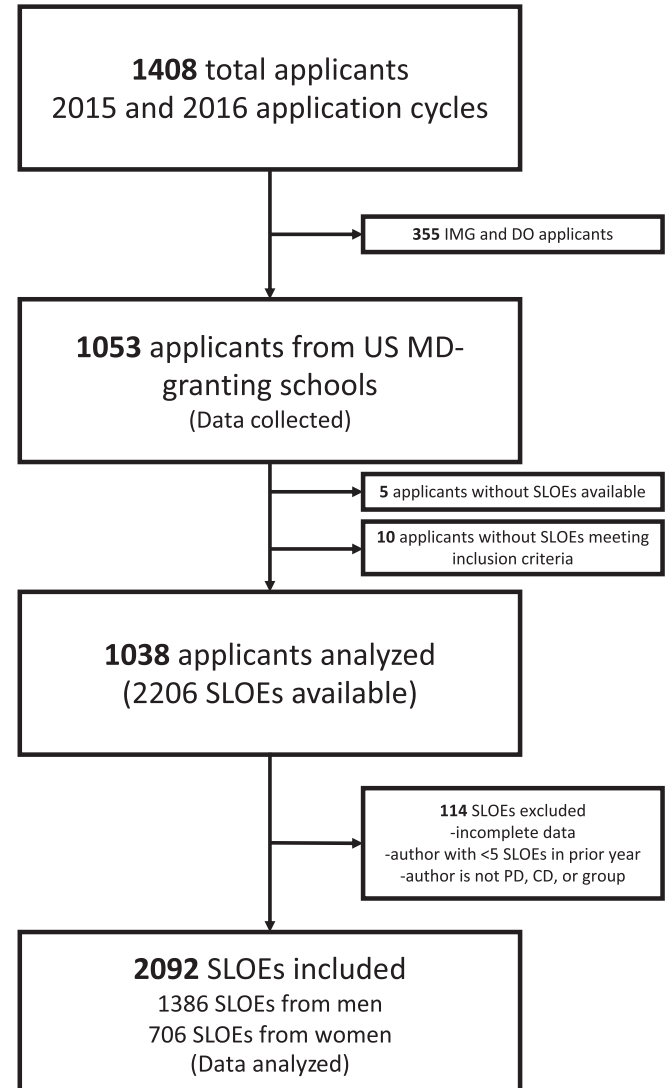


Figure 1. Flowsheet for study inclusion and exclusion. CD = clerkship director; DO = doctor of osteopathic medicine; IMG = international medical graduate; PD = program director; SLOE = standardized letter of evaluation.

five SLOEs in the previous year (Figure 1). Of the 2,092 included SLOEs, 1,386 (66%) SLOEs were from male applicants.

Students' medical schools were in the following regions: west ($n = 294$, 28%), midwest ($n = 236$, 23%), northeast ($n = 254$, 24%), and south ($n = 254$, 24%). The mean (\pm SD) step 1 score for men in the study was 232.0 (\pm 16.4) and was 228.2 (\pm 16.7) for women ($p < 0.001$). The mean (\pm SD) USMLE step 2 score for men in the study was 241.8 (\pm 14.6) and 243.7 (\pm 13.9) for women ($p = 0.035$).

Women achieved slightly lower, and thus better, CRSs, RLPSs, and CSs than men. The mean CRS for women was 2.27 and was 2.45 for men ($p < 0.001$). The mean RLPS for women was 2.32 and was 2.52 for men ($p < 0.001$). The mean CS was 4.59 for women and 4.97 for men ($p < 0.001$). Nineteen percent of women received a "top 10%" rating on the CRS ($n = 134$) and the RLPS ($n = 134$); 15% of men received a top 10% CRS ($n = 203$) and 14%

earned a top 10% RLPS ($n = 195$; $p < 0.001$). Figure 2 gives the distribution for each category rank in the SLOE by gender.

DISCUSSION

Differences in the assessment of men and women have been documented within undergraduate and graduate medical education.¹⁸⁻²⁸ Most notably and recently, Dayal et al.²⁰ studied the effect of gender on milestone assessments in EM residency training programs. Dayal and colleagues found that at the beginning of training, milestone assessments tend to be similar between men and women, with women achieving higher mean scores in many categories. At the end of training, however, men receive higher milestone assessments.²⁰

Our study was designed to observe if any gender differences exist at the earliest phase of EM training, the EM clerkship. We found that, from a large pool

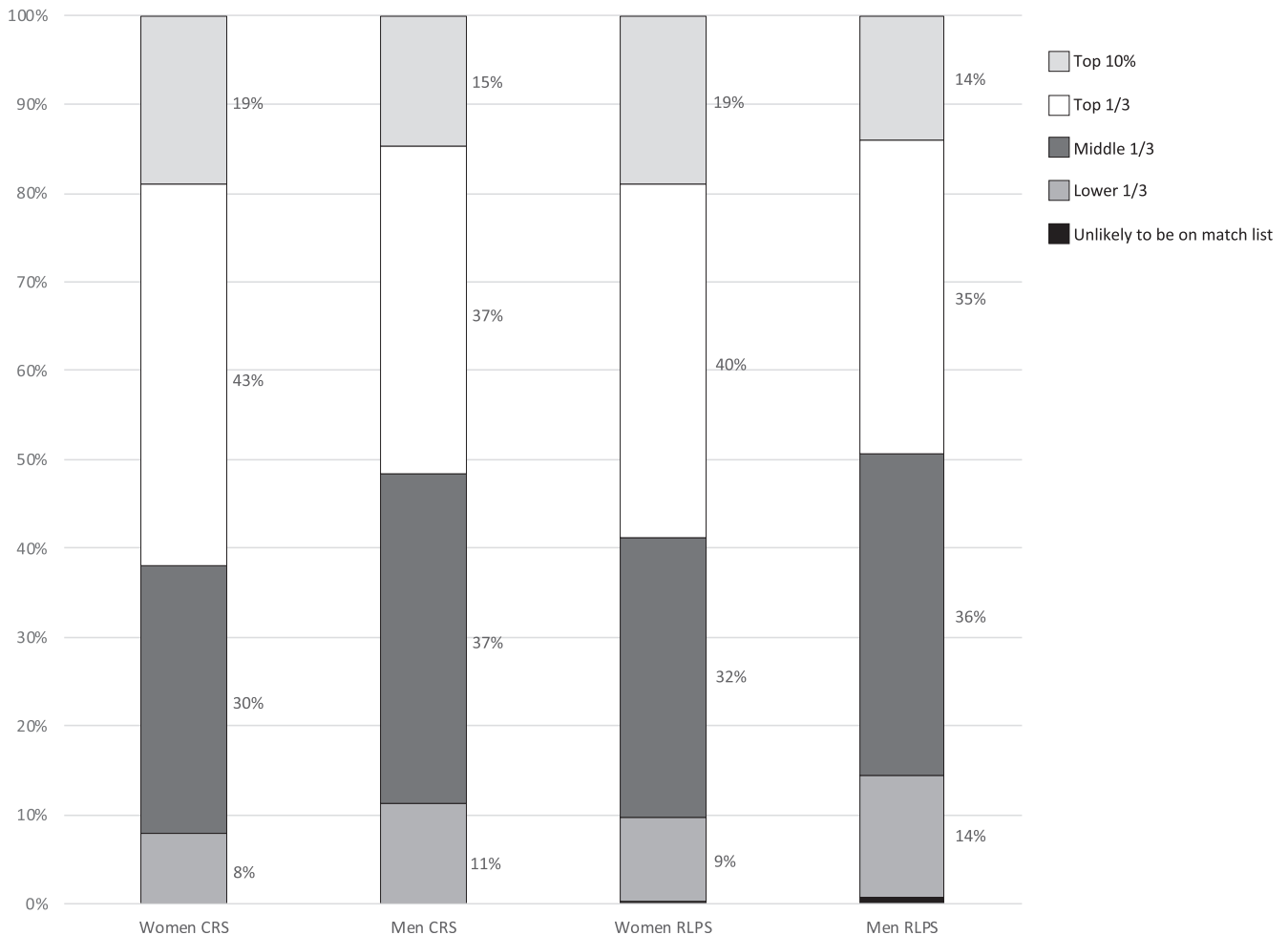


Figure 2. Distribution of SLOE CRS and RLPS for men and women. CRS = comparative rank score; RLPS = rank list position score; SLOE = standardized letter of evaluation

of EM applicants, female students are assessed somewhat more positively than male students in the global assessment on the SLOE. Although our study findings were statistically significant, the difference between male and female CSs (-0.38) was relatively small and corresponds to one quantile difference (i.e., top third vs. middle third in one category) for every three men.

It is unclear what is driving the discrepancies between gender and performance at each stage of training. Women may perform slightly better during the clerkship (our findings). Both genders perform similarly during intern year on milestone assessments.²⁰ Men may perform better during their final year of residency on milestone assessments.²⁰ It is possible that qualities important for the success of a senior resident are different than the qualities important for success of a fourth-year medical student. It is possible that a person's gender may offer qualities that predispose them to certain skills and attitudes. On the other hand, it is possible that one gender may be more negatively assessed than another gender when they exhibit certain qualities (e.g., assertiveness in women or uncertainty in men). Perhaps, in their role as subinterns, male medical students may suffer from gender bias and are perceived more negatively than their female counterparts. Alternatively, female medical students may perform better in this particular clinical environment. This could also be a function of different constructs being assessed in the SLOE versus milestones.

Haist et al.²⁶ found that women performed better than men on clinical examinations and suggested it may be due to their improved ability to deal with uncertainty, which is much higher in clinical years than basic science years. While men had higher USMLE Step 1 scores, women had higher USMLE Step 2 CK scores in our study. This is consistent with previous studies on the USMLE and gender.^{28,29} The USMLE Step 2 CK score difference could be responsible for the gender difference observed in our study, since the USMLE Step 2 CK tests clinical knowledge and its application. Ultimately, the cause for our observed gender differences EM clerkship assessment is unclear.

Moving forward, it would be interesting to observe longitudinal assessments of medical students from admission to residency. As we move toward competency-based assessment for undergraduate and graduate medical education, it is essential that we acknowledge the risk of gender and other biases.

LIMITATIONS

We only included applications to a single residency program; however, the regional distribution of medical schools was diverse and there were a large number of applicants to our program. Our female study subjects performed better on the USMLE Step 2 CK. While this is consistent with national trends,²⁸ our observed gender difference could be consistent with actual differences in clinical skills. We do not know if SLOE writers had USMLE Step 2 CK performance data available when assigning a RLPS. The leadership of our residency program is largely female and could have influenced the gender of applicants applying to our program. The gender distribution of residents during the study period was 52% female. Although our study findings were statistically significant, the difference between male and female CSs was less than one quantile difference.

CONCLUSIONS

The standardized letter of evaluation is an important part of a student's application to emergency medicine residency. Women receive slightly lower, and thus better, mean rank list position score, comparative rank score, and composite score when compared to men. Women receive a top 10% and top third rating with higher frequency than men. It is unclear what accounts for this difference.

References

1. Keim S, Rein J, Chisholm C, et al. A standardized letter of recommendation for residency application. *Acad Emerg Med* 1999;6:1141–6.
2. Martin DR, McNamara R. The CORD standardized letter of evaluation: have we achieved perfection or just a better understanding of our limitations? *J Grad Med Educ* 2014;6:353–4.
3. Love JN, Deiorio NM, Ronan-Bentle S, et al. Characterization of the council of emergency medicine residency directors' standardized letter of recommendation in 2011-2012. *Acad Emerg Med* 2013;20:926–32.
4. Love JN, Smith J, Weizberg M, et al. Council of Emergency Medicine Residency Directors' standardized letter of recommendation: the program director's perspective. *Acad Emerg Med* 2014;21:680–7.
5. Negaard M, Assimacopoulos E, Harland K, Van Heukelom J. Emergency medicine residency selection criteria: an update and comparison. *AEM Educ Train* 2018;2:146–53.

6. National Resident Matching Program, Data Release and Research Committee: Results of the 2018 NRMP Program Director Survey. National Resident Matching Program. Washington, DC. 2018. Available at: <https://www.nrmp.org/wp-content/uploads/2018/07/NRMP-2018-Program-Director-Survey-for-WWW.pdf>. Accessed June 10, 2019.
7. Council of Residency Directors in Emergency Medicine (CORD-EM) Advising Students Committee in EM (ASC-EM). Emergency Medicine Applying Guide. Available at: <https://www.cordem.org/globalassets/files/student-resources/em-applying-guide.pdf>. Accessed June 10, 2019.
8. CORD. The Standardized Letter of Evaluation (SLOE): Instructions for Authors. Available at: <https://www.cordem.org/esloe>. Accessed May 9, 2019.
9. Girzadas DV, Harwood RC, Dearie J, Garrett S. A comparison of standardized and narrative letters of recommendation. *Acad Emerg Med* 1998;5:1101–4.
10. Burkhardt JC, Stansfield RB, Vohra T, et al. Prognostic value of the multiple mini-interview for emergency medicine residency performance. *J Emerg Med* 2015;49:196–202.
11. Hayden SR, Hayden M, Gamst A. What characteristics of applicants to emergency medicine residency programs predict future success as an emergency medicine resident? *Acad Emerg Med* 2005;12:206–10.
12. Bhat R, Takenaka K, Levine B, et al. Predictors of a top performer during emergency medicine residency. *J Emerg Med* 2015;49:505–12.
13. Oyama LC, Kwon M, Fernandez JA, et al. Inaccuracy of the global assessment score in the emergency medicine standard letter of recommendation. *Acad Emerg Med* 2010;17(Suppl 2):38–41.
14. Hegarty CB, Lane DR, Love JN, et al. Council of Emergency Medicine Residency Directors standardized letter of recommendation writers' questionnaire. *J Grad Med Educ* 2014;6:301–6.
15. Grall KH, Hiller KM, Stoneking LR. Analysis of the evaluative components on the standard letter of recommendation (SLOR) in emergency medicine. *West J Emerg Med* 2014;15:419–23.
16. Jackson JS, Bond M, Love JN, Hegarty C. Emergency medicine standardized letter of evaluation (SLOE): findings from the new electronic SLOE format. *J Grad Med Educ* 2019:182–6.
17. Girzadas DV, Harwood RC, Davis N, Schulze L. Gender and the Council of Emergency Medicine Residency Directors standardized letter of recommendation. *Acad Emerg Med* 2004;11:988–91.
18. Ross DA, Boatright D, Nunez-Smith M, Jordan A, Chekroud A, Moore EZ. Differences in words used to describe racial and gender groups in medical student performance evaluations. *PLoS ONE* 2017;12:e0181659.
19. 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;177:651–7.
20. Rojek AE, Khanna R, Yim JW, et al. Differences in narrative language in evaluations of medical students by gender and under-represented minority status. *J Gen Intern Med* 2019;34:684–91.
21. Rand VE, Hudes ES, Browner WS, Wachter RM, Avins AL. Effect of evaluator and resident gender on the American Board of Internal Medicine evaluation scores. *J Gen Intern Med* 1998;13:670–4.
22. Klein R, Julian KA, Snyder ED, et al. Gender bias in resident assessment in graduate medical education: review of the literature. *J Gen Intern Med* 2019;34:712–9.
23. Thackeray EW, Halvorsen AJ, Ficalora RD, Engstler GJ, McDonald FS, Oxentenko AS. The effects of gender and age on evaluation of trainees and faculty in gastroenterology. *Am J Gastroenterol* 2012;107:1610–4.
24. 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;9:577–85.
25. Haist SA, Wilson JF, Elam CL, Blue AV, Fosson SE. The effect of gender and age on medical school performance: an important interaction. *Adv Heal Sci Educ* 2000;5:197–205.
26. Krueger PM. Do women medical students outperform men in obstetrics and gynecology? *Acad Med* 2006;73:101–2.
27. Cuddy MM, Swanson DB, Clauser BE. A multilevel analysis of the relationships between examinee gender and United States Medical Licensing Exam (USMLE) step 2 CK content area performance. *Acad Med* 2007;82(10 Suppl):89–93.
28. Cuddy MM, Swanson DB, Clauser BE. A multilevel analysis of examinee gender and USMLE Step 1 performance. *Acad Med* 2008;83:S58–62.
29. Li S, Fant A, McCarthy D, Miller D, Craig J, Kontrick A. Gender differences in language of standardized letter of evaluation narratives for emergency medicine residency applicants. *AEM Educ Train* 2017;1:334–9.
30. Association of American Medical Colleges and Electronic Residency Application Services. Emergency Medicine. Available at: <https://www.aamc.org/download/358770/data/emergencymed.pdf>. Accessed June 9, 2019.
31. Association of American Medical Colleges (AAMC). ACGME Residents and Fellows by Sex and Specialty, 2015. Available at: <https://www.aamc.org/data/workforce/reports/458766/2-2-chart.html>. Accessed May 6, 2019.