

UC Irvine

Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health

Title

A Novel Method for Blinding Reviewers to Gender for the Purposes of Gender Bias Research

Permalink

<https://escholarship.org/uc/item/09j844j4>

Journal

Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health, 20(4.1)

ISSN

1936-900X

Authors

Pallaci, M
Aluisio, A
See, A
et al.

Publication Date

2019

Copyright Information

Copyright 2019 by the author(s). This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

emergency medicine (EM) residency application. Most programs require at least one Standardized Letter of Evaluation (eSLOE) from an away institution, which necessitates visiting rotations. (An eSLOE is filed electronically on the eSLOE website created by the Council of Residency Directors in Emergency Medicine.) However, there is little information regarding the geographic distribution of EM clerkship positions compared to ACGME-approved postgraduate year (PGY)-1 positions.

Objectives: Our goal was to examine the geographic distribution of EM clerkship spots to ACGME-approved PGY-1 residency positions. Through a geographic distribution, state and regional imbalances in academic EM training can be identified. Our hypothesis was that states with a large number of EM graduate medical education (GME) spots would also have a high capacity of EM undergraduate medical education (UME) clerkship spots.

Methods: We performed an observational, retrospective analysis of all EM eSLOEs written during the 2016-2017 residency application cycle, categorizing authorship by institution and state. All ACGME-approved EM residency programs in the United States, registered in the 2017 National Resident Matching Program were also categorized by state. UME to GME ratios were calculated for each State and the geographic distribution was determined. We also identified the top 25 institutions with the largest EM UME volume.

Results: In 2016-2017, 6715 eSLOEs were submitted on 3138 unique EM applicants, written at 222 different institutions. Ratios of UME to GME positions varied from 1.6 in Connecticut to 21 in Vermont, with a nationwide average of 3.3. The top 25 institutions authored 1660 eSLOEs, accounting for 25% of the total eSLOEs.

Conclusion: There were large differences in the geographic distribution of UME and GME EM positions in 2016-2017. These data identify states and institutions with large EM clerkship capacities, which can enhance EM student advising for visiting rotations and geographic residency goals. EM subspecialty elective rotations and institutions with non-group eSLOE authorship are potential confounding limitations to this research.

State	Number of GME Spots	Number of UME spots	Ratio of UME/GME	State	Number of GME Spots	Number of UME spots	Ratio of UME/GME
Vermont	0	21	21	Puerto Rico	17	52	3.1
Oklahoma	8	78	9.5	Arkansas	10	31	3.1
Alabama	10	67	6.7	Michigan	137	425	3.1
Iowa	9	57	6.3	Indiana	21	62	3
District of Columbia	21	117	5.6	Hawaii	0	3	3
Nevada	8	45	5.6	Tennessee	34	102	3
New Jersey	60	264	4.4	Massachusetts	68	188	2.8
Louisiana	32	136	4.3	Mississippi	13	37	2.8
Georgia	25	107	4.3	Pennsylvania	154	413	2.7
Dalaware	12	50	4.2	Washington	12	31	2.6
New Mexico	12	50	4.2	South Carolina	41	108	2.6
Virginia	41	171	4.2	Utah	11	29	2.6
Maine	10	41	4.1	Arizona	35	87	2.5
Rhode Island	15	62	4.1	Kansas	8	20	2.5
Maryland	25	100	4	North Carolina	61	154	2.5
Minnesota	30	117	3.9	Missouri	39	99	2.5
West Virginia	10	39	3.9	Nebraska	10	20	2
Wisconsin	22	86	3.9	South Dakota	0	2	2
Colorado	17	67	3.9	Florida	118	221	1.9
Ohio	97	371	3.8	Connecticut	33	54	1.6
New York	297	1063	3.6	Alaska	0	0	0
Oregon	11	37	3.4	Idaho	0	0	0
Texas	121	399	3.3	Montana	0	0	0
Illinois	120	383	3.2	New Hampshire	6	0	0
Kentucky	22	69	3.1	North Dakota	0	0	0
California	184	579	3.1	Wyoming	0	0	0
				Total	2047	6712	3.3

27 A Novel Method for Blinding Reviewers to Gender for the Purposes of Gender Bias Research

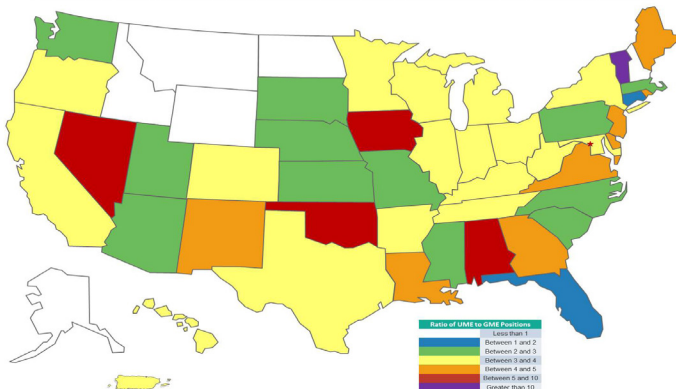
Pallaci M, Aluisio A, See A, Weinstock M, Beck-Esmay J, Frye A, Riddell J / Adena Health System Emergency Medicine Residency, Chillicothe, Ohio; Warren Alpert School of Medicine, Brown University, Providence, Rhode Island; Mount Sinai St. Luke's-Mount Sinai West, New York, New York; Keck School of Medicine of the University of Southern California, Los Angeles, California

Background: Research has shown a gender gap on ACGME Milestone evaluations, with the largest differences in procedural competency. There are currently no established methods by which researchers can blind reviewers to gender to evaluate for bias.

Objectives: This study evaluated whether a novel method could effectively blind evaluators to the gender of trainees performing simulated procedures. Since correct gender identification would be 50% based on chance, it was hypothesized that blinding would be effective if the reviewers did not correctly identify proceduralists' genders in $\geq 50\%$ of assessments.

Methods: After removing all jewelry from their hands, proceduralists were gowned, double-gloved, and filmed by a professional videographer while performing simulated procedures. Only their double-gloved hands, gowned forearms and lower torsos were visible in the videos. Five residents (two male and three female) performed three procedures each (lumbar puncture, chest tube, and central line placement), yielding 15 videos. Short video clips (30-45 seconds) were scored by seven graduate medical educators on a Likert scale (1 = definitely male, 3 = likely male, 5 = can't tell, 7 = likely female, 9 = definitely female). A response concordant with proceduralist gender with a confidence level of likely or higher (1-3 for males, 7-9 for females) was considered correct

Distribution of Emergency Medicine Undergraduate Medical Education Positions vs. Graduate Medical Education Positions in the United States



gender identification. Gender was considered to have not been correctly identified if the reviewer was unsure (score 4-6) or if the score was discordant with proceduralist gender. Reviewer scores were described, and we used one sample equality of proportions to assess significances against the null of >50% correct gender identification.

Results: Of 105 total responses, 56 (53.3%) expressed confidence in the gender of the proceduralist (1-3 or 7-9). Across all reviewers and procedures, 30.8% (95% confidence interval, 22.7-39.9%) correctly identified the proceduralist's gender. This proportion was statistically non-significant as compared to the null of >50% correct gender identification ($p = 1.00$). The same non-significant differences were maintained when the data were stratified by each procedure assessed.

Conclusions: Our method effectively blinded reviewers to the gender of the proceduralist and represents an innovative approach to perform needed research pertaining to gender bias in education.

Video Sample	Code	Reviewers							Mean score per individual operator	Operator
		Reviewer1	Reviewer2	Reviewer3	Reviewer4	Reviewer5	Reviewer6	Reviewer7		
1	9	3	5	9	6	5	7	9	6.3	F1
2	1	5	3	1	3	1	3	1	2.4	M1
3	1	5	7	4	7	7	2	5	5.3	M2
4	1	5	5	6	2	5	2	5	4.3	M2
5	1	7	6	1	8	6	2	2	4.6	M1
6	9	5	7	9	4	5	1	7	5.4	F1
7	9	4	5	9	7	5	6	6	6.0	F2
8	9	7	5	9	5	5	7	6	6.3	F2
9	9	7	4	1	7	5	2	4	4.3	F1
10	1	7	7	9	5	5	2	3	5.4	M2
11	1	3	7	1	3	5	5	4	4.0	M1
12	9	6	6	9	7	5	7	6	6.6	F3
13	9	6	7	3	7	5	3	1	4.6	F2
14	9	6	5	1	5	5	3	2	3.9	F3
15	9	5	5	9	5	5	1	5	5.0	F3
Percent Correctly Identified	0%	0%	60%	0%	7%	0%	13%	11%		

Individual Mean for Operators	Reviewer Ratings	Operator Code
F1 5.3	1= Definitely Male	1= Male
M1 3.7	3= Likely Male	9= Female
M2 5.0	5= unable to determine male or female	
F2 5.6	7= Likely Female	
F3 5.1	9= Definitely Female	

competencies or subsequent resident performance/success. While the SVI is currently being used without added cost to the applicant, it is projected that this will not be the case in the future. Thus, this additional student and training program expense should provide residency programs with enough objective data to make it worth the cost; these data do not yet exist.

Objectives: This study evaluated if a correlation existed between an applicant's SVI score and his or her position within a program's rank list during the 2018 National Resident Matching Program (NRMP) cycle. We hypothesized that the SVI score would not correlate with an applicant's place on an emergency medicine (EM) rank list.

Methods: We performed a prospective, survey-based study to evaluate correlation between an applicant's SVI score, and rank location by EM programs during the 2018 NRMP match. Inclusion criterion was defined as any ACGME-accredited EM program that participated in the NRMP match. Collected data included program demographics and de-identified applicant information, including rank and SVI score.

Results: Five programs participated in the survey study. Four programs were three years in length and one was four years. Three programs reported an academic setting, whereas the other two were community programs. On average, participating programs interviewed 175 applicants out of an average of 1003 applications received during the 2018 match cycle. We found no correlation between SVI score and position on a program's rank list (coefficient = -0.121, p-value = 0.076).

Conclusion: The results of this study revealed no correlation between a student's SVI score and his or her position on a program's rank list. Based on this small study, we question the utility of the SVI score in aiding programs during the recruitment process. Larger studies will be needed in the future to validate these results.

28 Standard Video Interview Scores and Applicant Position on Residency Program List: A Correlation Study

Bowers K, Comp G, Kalnow A, Casey J, Fraser W, Lloyd C, Little A / Doctors Hospital Ohio Health, Columbus, Ohio

Background: The Standard Video Interview (SVI) is an interview tool developed by the Association of American Medical Colleges to assess professional and interpersonal/communication skills in medical students applying for residency. The individual SVIs for each applicant are graded on 1-5 scale with the sum of all interviews used to create the applicants overall SVI score ranging from 6-30. The higher the score, the more the applicant demonstrated proficiency in managing difficult situations presented to them. There are conflicting opinions on the value and usefulness of the SVI, as scoring does not yet have a known correlation with ACGME

