

Diversity, Equity, Inclusion, and Justice

Ethnic and Racial Differences in Ratings in the Medical Student Standardized Letters of Evaluation (SLOE)

Al'ai Alvarez¹, MD
 Alexandra Mannix, MD
 Dayle Davenport, MD
 Katarzyna Gore, MD
 Sara M. Krzyzaniak, MD
 Melissa Parsons, MD

Danielle T. Miller², MD, MEd
 Daniel Eraso, MD
 Sandra Monteiro³, PhD
 Teresa M. Chan⁴, MD, MHPE
 Michael Gottlieb⁵, MD

ABSTRACT

Background The Standardized Letter of Evaluation (SLOE) stratifies the assessment of emergency medicine (EM) bound medical applicants. However, bias in SLOE, particularly regarding race and ethnicity, is an underexplored area.

Objective This study aims to assess whether underrepresented in medicine (UIM) and non-UIM applicants are rated differently in SLOE components.

Methods This was a cross-section study of EM-bound applicants across 3 geographically distinct US training programs during the 2019-2020 application cycle. Using descriptive and regression analyses, we examine the differences between UIM applicants and non-UIM applicants for each of the SLOE components: 7 qualifications of an EM physician (7QEM), global assessment (GA) rating, and projected rank list (RL) position.

Results Out of a combined total of 3759, 2002 (53.3%) unique EM-bound applicants were included. UIM applicants had lower ratings for each of the 7QEM questions, GA, and RL positions. Compared to non-UIM applicants, only some of the 7QEM components: "Work ethic and ability to assume responsibility," "Ability to work in a team, and "Ability to communicate a caring nature," were associated with their SLOE. "Commitment to EM" correlated more with GA for UIM than for non-UIM applicants.

Conclusions This study shows a difference in SLOE rating, with UIM applicants receiving lower ratings than non-UIM applicants.

Introduction

Underrepresented in medicine (UIM) students face multiple systemic barriers, including bias and discrimination.^{1,2} The Association of American Medical Colleges (AAMC) describes UIM as "racial and ethnic populations that are underrepresented in the medical profession relative to their numbers in the general population."³ UIM students experience discrimination, stereotyping, and racial disparities in clinical grading and evaluation.⁴⁻⁸

The Standardized Letter of Evaluation (SLOE) provides evaluative assessments of students⁹ critical to residency screening and selection,^{10,11} and is a composite of a grade, global assessment (GA), predicted rank list position (RL), and assessment of the 7 qualifications of an emergency medicine (EM) physician (7QEM). A SLOE template is provided as

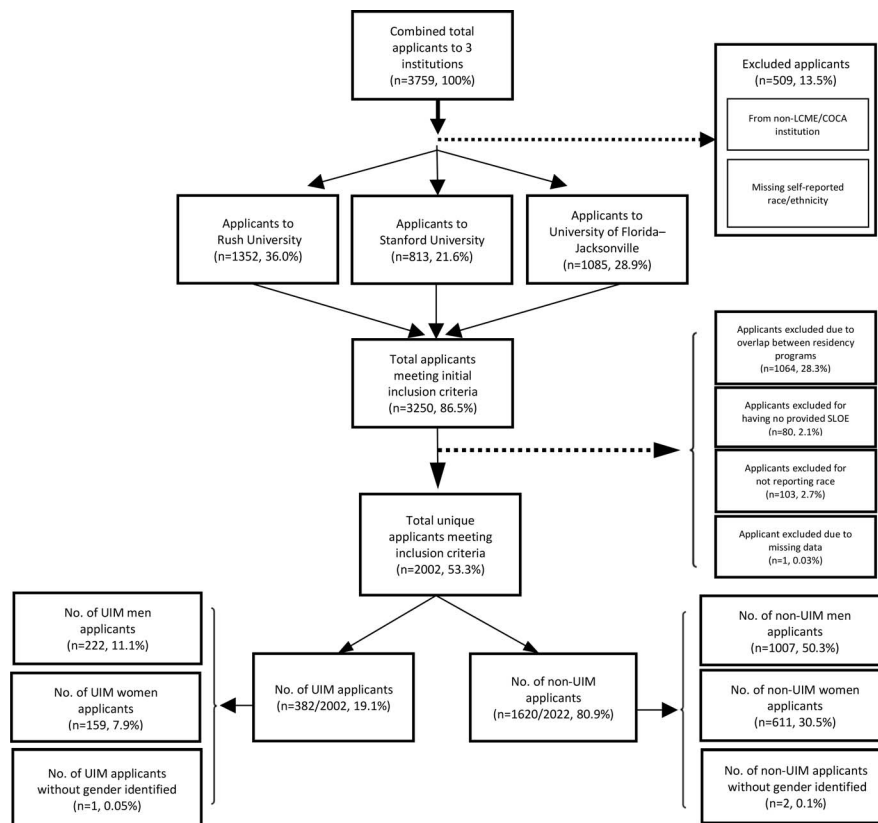
online supplementary data. As United States Medical Licensing Examination (USMLE) Step 1 transitions to pass/fail, the SLOE will likely carry greater weight in applicant selection.^{12,13} Prior research illustrates the presence of gender differences,^{14,15} but there are no data on racial or ethnic differences in SLOEs.

Methods

We performed a multi-institutional, cross-sectional, convenience sample study of SLOEs among 3 US EM residency programs (Rush University, Stanford University, and University of Florida–Jacksonville) through the Electronic Residency Application Service (ERAS) during the 2019-2020 application cycle. The programs represented distinct program lengths, types, and geography. The academic year selected for study was before COVID-19 and USMLE Step 1 becoming pass/fail. We defined UIM using the AAMC initial definition of "racial groups of Black, Mexican-American, mainland Puerto Rican, and Native American (American Indian and natives of Alaska and Hawaii)" because

DOI: <http://dx.doi.org/10.4300/JGME-D-21-01174.1>

Editor's Note: The online version of this article contains the SLOE template and further analysis from the study.



FIGURE

Inclusion Criteria of Total Applications Across all 3 Institutions With Breakdown of Applicants by Self-Identified Race and Gender

Abbreviations: LCME, Liaison Committee on Medical Education; COCA, Commission on Osteopathic College Accreditation; SLOE, Standardized Letters of Evaluation; UIM, underrepresented in medicine.

ERAS allows filtering for these self-reported variables.³ Only a single application was reviewed when an applicant applied to multiple programs. Our theoretical orientation is post-positivist, and the theoretical framework is aligned with a post-colonial lens to examine the effect of race on modern structures.¹⁶

We included all applicants from the Liaison Committee for Medical Education or Commission on Osteopathic College Accreditation accredited medical schools who applied to at least one of the institutions.

Abstractors from each institution used a pre-piloted standardized data abstraction tool to collect the following data from ERAS: AAMC number, self-identified gender and race/ethnicity, and medical school. For each SLOE, we collected the rating for all 7QEM questions, GA, and predicted RL position. As with prior literature, order values were assigned for GA, RL, and 7QEM.^{17,18} The first 5QEM used the following anchors: (1) “Below peers,” (2) “At level of peers,” and (3) “above peers.” The sixth QEM used (1) “More than peers,” (2) “Same as peers,” and (3) “Less than peers.” The seventh QEM used (1)

“Good,” (2) “Excellent,” and (3) “Outstanding.” For GA and RL, students are assessed in comparison to other applicants as top 10% (4), top third (3), middle third (2), or lower third (1). With each applicant included in the study having a single averaged rating for each SLOE component, the data were treated as continuous variables.^{17,18}

We used a repeated measures analysis of variance to examine the dependent variable of 7QEM ratings with one independent variable of UIM status to determine if mean ratings differed between groups. All analyses were computed using SPSS 26 (IBM Corp, Armonk, NY). Data were recorded and coded using a Microsoft Excel spreadsheet. Further discussion of the analytic approach is included in the online supplementary data.

This study was granted an exemption from all 3 Institutional Review Boards.

Results

Out of 3759 applicants, 3250 (86.5%) met the initial inclusion criteria. Of these, 1248 (38.4%) applicants

TABLE
Multiple Linear Regression Analysis of 7 Qualities of an EM Physician (7QEM) Ratings as They Correlate With Global Assessment (GA) and Rank List (RL)

7 Qualifications for EM (7QEM)	Candidate Cohort Mean Out of 3 Institutions (SD)		Regression Model for Global Assessment (GA)		Regression Model for Rank List (RL)	
	UIM	non-UIM	UIM	non-UIM	UIM	non-UIM
	Standardized Beta Coefficient (95% CI, t Test, P Value)		Standardized Beta Coefficient (95% CI, t Test, P Value)		Standardized Beta Coefficient (95% CI, t Test, P Value)	
Commitment to emergency medicine	2.48 (0.41)	2.45 (0.43)	0.12 (0.09-0.32, 3.39, P<.001)	0.06 (0.06-0.16, 3.98, P<.001)	0.01 (0-0.16, 0.37, P=.71)	0.05 (0.03-0.15, 3.04, P<.01)
	$\eta^2=0.000$					
Work ethic, willingness to assume responsibility	2.62 (0.44)	2.66 (0.39)	0.03 (0-0.17, 0.59, P=.56)	0.12 (0.14-0.26, 6.21, P<.001)	0.06 (0-0.24, 1.30, P=.19)	0.12 (0.16-0.29, 6.4, P<.001)
	$\eta^2=0.002$					
Ability to develop and justify an appropriate differential and a cohesive treatment plan	2.18 (0.52)	2.31 (0.47)	0.29 (0.28-0.51, 6.7, P<.001)	0.26 (0.32-0.43, 13.00, P<.001)	0.21 (0.18-0.43, 4.71, P<.001)	0.18 (0.21-0.34, 8.7, P<.001)
	$\eta^2=0.011$					
Ability to work with a team	2.55 (0.42)	2.57 (0.42)	0.05 (0-0.22, 1.29, P=.20)	0.11 (0.11-0.24, 5.45, P<.001)	0.12 (0.08-0.37, 2.97, P<.01)	0.15 (0.19-0.33, 7.39, P<.001)
	$\eta^2=0.000$					
Ability to communicate a caring nature to patients	2.57 (0.40)	2.52 (0.41)	0.06 (0-0.23, 1.49, P<.14)	0.06 (0.04-0.15, 3.18, P<.002)	0.11 (0.08-0.36, 3.02, P<.01)	0.08 (0.07-0.20, 4.17, P<.001)
	$\eta^2=0.002$					
How much guidance do you predict this applicant will need during residency?	2.10 (0.53)	2.27 (0.47)	0.25 (0.20-0.46, 4.98, P<.001)	0.21 (0.23-0.36, 9.67, P<.001)	0.28 (0.25-0.54, 5.42, P<.001)	0.22 (0.27-0.41, 9.55, P<.001)
	$\eta^2=0.020$					
Given the necessary guidance, what is your prediction of success for the applicants?	2.25 (0.50)	2.35 (0.49)	0.21 (0.17-0.43, 4.44, P<.001)	0.22 (0.25-0.37, 9.67, P<.001)	0.23 (0.20-0.49, 4.71, P<.001)	0.23 (0.28-0.41, 9.96, P<.001)
	$\eta^2=0.007$					

Abbreviations: EM, emergency medicine; UIM, underrepresented in medicine.

were excluded. Exclusion criteria and demographics of all applicants meeting inclusion criteria can be found in the FIGURE. Our included applicants represented 58.8% (2002 of 3405) of all EM applicants for the 2019-2020 application cycle ERAS data.¹⁹ The included applicants contributed 5433 SLOEs to the data set, with 4717 SLOEs meeting inclusion criteria. A total of 716 SLOEs were excluded: 60 subspecialty SLOEs, 118 with incomplete data, 157 not written by program leadership, 425 written by a letter writer who wrote <10 SLOEs the previous year, and 44 after meeting multiple exclusion criteria. Of the 4717 SLOEs included, 891 (18.9%) were from UIM applicants.

UIM applicants received lower average 7QEM ratings (2.39 vs 2.45, $\eta^2=0.01$), mean GA ratings (2.40 vs 2.59, $\eta^2=0.01$), and mean RL rankings (2.42 vs 2.59, $\eta^2=0.01$). Consistent with prior research,^{17,18} GA and RL were converted into anchors with differences in percentages found to be statistically significant. The TABLE displays the mean rating for each QEM and effect size, reported as partial eta squared (η^2).

While all 7QEM ratings had significant correlations with GA, the linear regression model revealed that only ratings on “Commitment to EM,” “Ability to develop a differential,” “Guidance needed,” and “Prediction of success” were associated with GA for UIM applicants. Additional linear regression model highlighted that while all 7QEM ratings were significantly correlated with RL, only ratings on “Ability to develop a differential,” “Ability to work with a team,” “Ability to communicate a caring nature,” “Guidance needed,” and “Prediction of success” are critically associated with the RL position for UIM applicants (TABLE).

Discussion

To our knowledge, this is the first study to specifically assess for differences in SLOE scoring for UIM and non-UIM in GA, RL, and all 7QEM questions. We found a difference in SLOE ratings, with UIM applicants receiving lower ratings than non-UIM applicants. The effect sizes are small yet consistent in all findings and may represent systematic bias. Findings noted add to growing literature recognizing UIM students’ experience of pervasive bias and discrimination in medical education.⁴⁻⁸ With USMLE Step 1 transitioning to pass/fail,^{12,13} the SLOE may be vulnerable to bias and should be examined further.

This study is subject to the limitations inherent to cross-sectional research. The data are limited to select EM programs and may not reflect all EM SLOEs. Additionally, we defined UIM using the AAMC definition,³ although bias may likewise have occurred

against other races not self-identifying as UIM. While we identified differences in the SLOE domain ratings of UIM vs non-UIM applicants, the relevance of these findings remains unclear as no data exist to date on how a specific rating will affect residency ranking and Match success. In addition, a significant limitation is that we did not correct for other application variables. Many other potential sources of bias beyond structural racism may contribute to different ratings in the SLOE. This study was not designed to elucidate the causal factors. Finally, our use of the multiple linear regression analysis is exploratory and should be interpreted with caution.

Bias exists in different aspects of the residency application. Our work highlights differential ratings in the SLOE for UIM vs non-UIM applicants. Future work should investigate how these differences impact the ranking of applicants. Additionally, more work is needed to compare the SLOE with other objective evaluative tools concerning racial and ethnic equity in grading across different specialties. Finally, future studies could assess the impact of diversity training for faculty on SLOE scoring.

Conclusions

Our study evaluated the relationship between race/ethnicity and SLOE components. We found differences in the overall ratings of the 7QEM questions, the predictors of GA, and the anticipated RL position on SLOEs for UIM and non-UIM applicants.

References

1. Grimm LJ, Redmond RA, Campbell JC, Rosette AS. Gender and racial bias in radiology residency letters of recommendation. *J Am Coll Radiol*. 2020;17(1 Pt A):64-71. doi:10.1016/j.jacr.2019.08.008
2. Capers Q, Clinchot D, McDougle L, Greenwald AG. Implicit racial bias in medical school admissions. *Acad Med*. 2017;92(3):365-369. doi:10.1097/ACM.0000000000001388
3. Associate of American Medical Colleges. Underrepresented in Medicine Definition. Accessed November 4, 2021. <https://www.aamc.org/what-we-do/equity-diversity-inclusion/underrepresented-in-medicine>
4. Williams DR, Mohammed SA. Discrimination and racial disparities in health: evidence and needed research. *J Behav Med*. 2009;32(1):20-47. doi:10.1007/s10865-008-9185-0
5. Lewis TT, Cogburn CD, Williams DR. Self-reported experiences of discrimination and health: scientific advances, ongoing controversies, and emerging issues. *Annu Rev Clin Psychol*. 2015;11:407-440. doi:10.1146/annurev-clinpsy-032814-112728

6. Low D, Pollack SW, Liao ZC, et al. Racial/ethnic disparities in clinical grading in medical school. *Teach Learn Med.* 2019;31(5):487-496. doi:10.1080/10401334.2019.1597724
7. Lee KB, Vaishnavi SN, Lau SKM, Andriole DA, Jeffe DB. "Making the grade:" noncognitive predictors of medical students' clinical clerkship grades. *J Natl Med Assoc.* 2007;99(10):1138-1150.
8. 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(8):e0181659. doi:10.1371/journal.pone.0181659
9. Keim SM, Rein JA, Chisholm C, et al. A standardized letter of recommendation for residency application. *Acad Emerg Med.* 1999;6(11):1141-1146. doi:10.1111/j.1553-2712.1999.tb00117.x
10. 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(6):680-687. doi:10.1111/acem.12384
11. Negaard M, Assimacopoulos E, Harland K, Van Heukelom J. Emergency medicine residency selection criteria: an update and comparison. *AEM Educ Train.* 2018;2(2):146-153. doi:10.1002/aet2.10089
12. Neville AL, Smith BR, de Virgilio C. USMLE Step 1 scoring system change to pass/fail-an opportunity for change. *JAMA Surg.* 2020;155(12):1093-1094. doi:10.1001/jamasurg.2020.2836
13. Cangialosi PT, Chung BC, Thielhelm TP, Camarda ND, Eiger DS. Medical students' reflections on the recent changes to the USMLE Step Exams. *Acad Med.* 2021;96(3):343-348. doi:10.1097/ACM.0000000000003847
14. Miller DT, McCarthy DM, Fant AL, Li-Sauerwine S, Ali A, Kontrick AV. The Standardized Letter of Evaluation narrative: differences in language use by gender. *West J Emerg Med.* 2019;20(6):948-956. doi:10.5811/westjem.2019.9.44307
15. Li S, Fant AL, McCarthy DM, 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(4):334-339. doi:10.1002/aet2.10057
16. Varpio L, Paradis E, Uijtdehaage S, Young M. The distinctions between theory, theoretical framework, and conceptual framework. *Acad Med.* 2020;95(7):989-994. doi:10.1097/ACM.0000000000003075
17. Mannix A, Monteiro S, Miller D, et al. Gender differences in emergency medicine Standardized Letters of Evaluation. *AEM Educ Train.* 2022;6(2):e10740. doi:10.1002/aet2.10740
18. Miller DT, Krzyzaniak S, Mannix A, et al. The Standardized Letter of Evaluation in emergency medicine: are the qualifications useful? *AEM Educ Train.* 2021;5(3):e10607. doi:10.1002/aet2.10607
19. Association of American Medical Colleges. ERAS Statistics. Accessed March 7, 2021. <https://www.aamc.org/data-reports/interactive-data/eras-statistics-data>



Al'ai Alvarez, MD, is Clinical Associate Professor and Director of Well-Being, Department of Emergency Medicine, Stanford University; **Alexandra Mannix, MD**, is Assistant Professor and Assistant Program Director, Department of Emergency Medicine, University of Florida College of Medicine-Jacksonville; **Dayle Davenport, MD**, is Associate Professor, Department of Emergency Medicine, Rush University Medical Center and Assistant Dean, Diversity and Inclusion, Rush Medical College; **Katarzyna Gore, MD**, is Associate Professor and Assistant Program Director, Department of Emergency Medicine, Rush University Medical Center; **Sara M. Krzyzaniak, MD**, is Clinical Associate Professor and Program Director, Department of Emergency Medicine, Stanford University; **Melissa Parsons, MD**, is Associate Professor and Assistant Program Director, Department of Emergency Medicine, University of Florida College of Medicine-Jacksonville; **Danielle T. Miller, MD, MEd**, is Assistant Professor, Department of Emergency Medicine, University of Colorado School of Medicine; **Daniel Eraso, MD**, is Assistant Professor, Department of Emergency Medicine, University of Florida College of Medicine-Jacksonville; **Sandra Monteiro, PhD**, is Associate Professor, Department of Medicine, McMaster University, Ontario, Canada; **Teresa M. Chan, MD, MHPE**, is Associate Professor, Department of Medicine (Division of Emergency Medicine; Division of Education & Innovation), and Associate Dean, Continuing Professional Development, Faculty of Health Sciences, McMaster University, Ontario, Canada; and **Michael Gottlieb, MD**, is Associate Professor and Ultrasound Director, Department of Emergency Medicine, Rush University Medical Center.

The authors would like to thank Dr. CJ Foote for his statistical analysis support on the preliminary data of this manuscript. We also thank Dr. Stefanie Sebok-Syer for her expertise and critical reading of the manuscript.

Funding: Dr. Gottlieb reports funding from the Centers for Disease Control and Prevention, Emergency Medicine Foundation, and Society for Academic Emergency Medicine.

Conflict of interest: The authors declare they have no competing interests.

This work was previously presented at CORDAA22, San Diego, CA, March 28, 2022; Society for Academic Emergency Medicine Western Regional Meeting, Stanford, CA, April 1, 2022; Society for Academic Emergency Medicine New England Regional Meeting, Providence, RI, April 6, 2022; and SAEM22, New Orleans, LA, May 12, 2022.

Corresponding author: Al'ai Alvarez, MD, Stanford University, al.ai.alvarez@stanford.edu, Twitter @alvarezzy

Received December 14, 2021; revisions received February 25, 2022, and August 12, 2022; accepted August 16, 2022.