

## Efforts to Reduce Bias in Clerkship Evaluations: A CERA Study

Nina Piazza, MD | Mechelle Sanders, PhD | Gerardo Moreno, MD, MS | Elizabeth Brown, MD, MPH

PRiMER. 2024;8:43.

Published: 8/5/2024 | DOI: 10.22454/PRiMER.2024.662375

### Abstract

**Introduction:** As the number of medical students who identify as underrepresented in medicine (URiM) increases, the disparities related to gender and URiM status persist. This study examines the current initiatives within family medicine clerkships to reduce bias in evaluations.

**Methods:** Our 10-item survey was included as a module in the 2022 Council of Academic Family Medicine Educational Research Alliance national survey of family medicine clerkship directors. Our survey questions asked about whether programs had strategies to reduce bias in student evaluations, antiracism initiatives, perceptions on effectiveness of the initiatives, and type and cadence of faculty development on evaluations for preceptors.

**Results:** The overall response rate for the survey was 59.12% (94/159); all respondents completed our module. Seventy percent said they had implemented strategies to reduce bias in evaluations, 60% felt these were effective, and 80% felt that reducing bias in evaluations was a priority. The majority, 89/91 (95%), indicated that their medical schools had a current social justice, diversity, or antiracism initiative. We identified a positive association between specific antibias medical school initiatives and clerkship directors undertaking practices to reduce bias in evaluations ( $P=.005$ ).

**Conclusions:** Most programs had implemented strategies to reduce bias and felt that doing so was a priority. Community-based preceptors were less likely to have faculty development around reducing bias compared to those in academics. Further improvements may need to prioritize including community preceptors in educational efforts to reduce bias.

## Introduction

Disparities and bias in evaluations and grading impact medical students' careers. Several studies on clerkship grades show disparities related to gender and underrepresented minority status, leading to a concern that bias in evaluation affects final grades.<sup>1-3</sup> Grading inequities are further worsened by the added pressures that students who are underrepresented in medicine (URiM) face, such as confronting stereotype threat and supervisors' implicit biases.<sup>4</sup> To students, the grading process can feel biased and inconsistent, leading to an unfavorable learning environment.<sup>4</sup>

Several studies have explored ways to reduce bias in the evaluation process, including using grading committees,<sup>5</sup> minimizing descriptive words that can lead to bias,<sup>1,6,7</sup> and using competency-based judgment of medical school performance.<sup>8-11</sup> However, a paucity of studies exist related to actual implementation of antibias evaluation strategies.

Our study examines the status of antibias initiatives within family medicine (FM) clerkships. We hypothesized that clerkships in medical schools that have current antiracism or diversity initiatives would have more strategies to reduce bias in evaluation and grading. Additionally, we hypothesized that being a female or URiM clerkship director (CD) would be positively associated with implementation strategies to reduce bias in evaluations.

## Methods

---

Our survey on antibias initiatives was included in the Council of Academic Family Medicine Educational Research Alliance (CERA) study of FM CDs. The data was collected from June 7, 2022, through July 8, 2022. The methodology of the survey has been previously described.<sup>12</sup> The survey is distributed annually to CDs of qualifying medical schools that are accredited by the Liaison Committee on Medical Education or the Committee on Accreditation of Canadian Medical Schools within the United States (n=148) and Canada (n=18), respectively. Our antibias survey included 10-questions about the use of strategies to reduce bias in medical student evaluations, antiracism initiatives, perceptions on effectiveness of the antibias initiatives, and type and cadence of faculty development on antibias evaluations. CD responses about faculty participation in development sessions and receiving evaluation instructions are based on their best estimates. The project was approved by the American Academy of Family Physicians Institutional Review Board.

All analyses were performed using STATA 13.1 (StataCorp). We used univariate statistics to describe CD and program characteristics. We used bivariate statistics to examine the relationships between X and Y. A *P*-level of <.05 defined statistical significance. We conducted bivariate analysis using  $\chi^2$  tests to calculate the associations between being at a medical school that conducted an analysis of grades and each of the following: implementing strategies to reduce bias in evaluations, effectiveness of an evaluation system designed to reduce bias, prioritizing strategies to reduce bias in evaluations, using standard behavior-focused anchors, annual faculty development for preceptors on how to write less biased descriptions, and attendance to faculty development of preceptors (community and faculty). We also examined the bivariate relationships between use of strategies to reduce bias with each gender, years graduated from residency, race, years as director, and protected time.

## Results

---

The response rate for the survey was 59.12% (94/159). All respondents completed our survey module. Respondents were primarily female (59.6%) and White (76.6%). Approximately 11% identified as URiM and, on average, had been a CD for 8 years (Table 1). The majority (70%, strongly agree or agree) had implemented strategies to reduce bias in evaluations and thought their bias-reducing strategies were at least somewhat effective (63%, somewhat or very effective; Table 2). Forty-two percent of community preceptors reported never attending faculty development on reducing bias in evaluations compared to 17% of the academic preceptors (Table 3). Table 2 shows that the majority of medical schools 89/91 (95%) had a current social justice, diversity, or antiracism initiative. One-third of the medical schools analyzed their bias in clerkship grading efforts while 15% were in the process of doing so.

In our bivariate analyses, CDs were more likely to implement strategies to reduce bias in evaluations if their medical school performed an analysis of grades compared to those that did not (90% vs 62%, *P*=.005).

Conducting a grade analysis was associated with an evaluation system designed to reduce bias (81% vs 56 %,  $P=.02$ ), prioritizing strategies to reduce bias in evaluations (100% vs 74 %,  $P=.002$ ), using standard behavior-focused anchors (81% vs 60%,  $P=.045$ ), and annual faculty development for preceptors on how to use less biased descriptors in their written evaluations (77% vs 34%,  $P=.045$ ).

Academic preceptors (71% vs 40%,  $P<.005$ ) and community preceptors (35% vs 11% ( $P<.005$ )) were more likely to attend faculty development on reducing bias in evaluations at least once a year if the medical school had performed an analysis of grades.

We found no statistically significant association between gender of CD nor race (White vs URiM) and use of various strategies to reduce bias in evaluations. CDs who agreed that they implemented strategies to reduce bias in evaluations had more recently graduated from residency (7.14 vs 10.44 years,  $P=.03$ ) and had a lower average number of years in their current director role (6.79 vs 10.24 years,  $P=.02$ ), as well as less protected time (30.2 vs 40.7,  $P=.04$ ).

## Discussion

---

In this study, we identified a positive correlation between medical school initiatives related to grade analysis and efforts to reduce bias in evaluations in FM clerkships. Reducing bias in evaluations was a priority for most FM CDs. Two-thirds of CDs believed their efforts were effective at reducing bias in evaluation. Faculty development on reducing bias in evaluation was less available to community preceptors compared to academic preceptors.

Our hypothesis that being a female or URiM CD would be associated with implementing more strategies to reduce bias was not supported by the data. We also found that CDs who were in the role for fewer years and had less protected time implemented strategies to reduce bias more often than those who had more protected time and had been a CD for a longer time.

The Alliance for Academic Internal Medicine recommends a comprehensive approach to reducing bias in medical education.<sup>10</sup> To our knowledge, an evidence-based faculty development program aimed at reducing bias in evaluation and grading does not currently exist. Such programs would be invaluable for reducing bias in evaluations for both academic and community preceptors.<sup>13</sup>

Our study's limitations included limited diversity in respondent demographics,<sup>14,15</sup> a modest response rate, recall bias of CDs with answers based on estimations, and inability to control for socially desirable answers. We did not ask about specific biases CDs were addressing, and some terms were not clearly defined for the respondents, such as "standard, behavior-focused anchors" and "antiracism and diversity initiatives." Also, given that the evaluations were presumed to be FM-specific, the results are not generalizable across all types of clerkships, and variation is likely across institutions. To the best of our knowledge, our study is among the first that explores FM CDs' initiatives to reduce bias in medical school evaluations. The results show that efforts can be impactful when medical schools have specific initiatives to analyze bias in clerkship grades and when academic faculty can more easily access training about bias reduction in evaluations. As strategies to address bias in evaluation are implemented, assessing for their intended effect is important. Future studies should evaluate long-term outcomes.

## Tables and Figures

---

**Table 1. Characteristics of Family Medicine Clerkship Directors\***

<b>Clerkship director characteristics</b>	<b>n (%)</b>
<b>Gender</b>	
Male	36 (38.3)
Female	56 (59.6)
Did not disclose	2 (2.1)
Genderqueer/gender nonconforming	0
Nonbinary	0
<b>Race/ethnicity</b>	
Hispanic/Latino/a/Spanish	0
American Indian/Alaska Native/Indigenous	0
Asian	13 (13.8)
Black or African American	4 (4.3)
Native Hawaiian or Pacific Islander	0
White	72 (76.6)
Middle Eastern/North African	1 (1.1)
Choose not to disclose	3 (3.2)
Underrepresented in medicine	10 (10.6)
Average number of years as clerkship director	7.6 (SD 6.0)
Physician clerkship director	93 (98.9)
<b>Years since residency</b>	
0-10	22 (24.7)
11-20	33 (37.0)
21-30	18 (20.2)
More than 30	14 (15.7)
Missing	2 (2.2)
<b>Percentage protected time for clerkship director</b>	
Less than 20%	12 (12.7)
20% -29%	29 (30.9)
30%-39%	24 (25.5)
40%-49%	12 (12.8)
50%-59%	10 (10.6)
60% or more	6 (6.4)
Missing	1 (1.1)
<b>Family medicine clerkship characteristics</b>	
Average number of evaluations a student receives	3.0 (SD 1.4)
Public medical school	65 (69.2)
<b>Number of medical students per year</b>	
Less than 100	23 (24.5)
100-199	50 (53.2)
200-299	19 (20.2)
300 or more	2 (2.1)
<b>Clerkship location/region**</b>	
Northeast	18 (19.2)
South Atlantic	20 (21.3)
South	14 (14.9)
Midwest	18 (19.1)

West	11 (11.7)
Canada	13 (13.8)

\*Table reports univariate statistics. Please see Results section for discussion of bivariate results.

\*\*Northeast=New England (NH, MA, ME, VT, RI, or CT) and Middle Atlantic (NY, PA, or NJ); West=Pacific (WA, OR, CA, AK, or HI) and Mountain (MT, ID, WY, NV, UT, AZ, CO, or NM); Midwest=East North Central (WI, MI, OH, IN, or IL) and West North Central (ND, MN, SD, IA, NE, KS, or MO); South=West South Central (OK, AR, LA, or TX) and East South Central (KY, TN, MS, or AL); and South Atlantic=FL, GA, SC, NC, VA, DC, WV, DE, PR, or MD.

Abbreviation: SD, standard deviation

**Table 2. Family Medicine Clerkship Evaluation Anchors, Bias Reduction, and Medical School Initiatives\***

Survey question	n (%)
<b>Medical school has current social justice, antiracism initiative, or diversity initiative.</b>	
Strongly agree	58 (61.7)
Agree	31 (33.0)
Neutral	2 (2.2)
Disagree	0
Strongly disagree	0
Missing	3 (3.2)
<b>Medical school clerkship analysis of grading to evaluate bias</b>	
Yes, does this regularly	13 (13.8)
Yes, have done this once	18 (19.2)
Currently working on this	14 (14.9)
No, but planning to do so	12 (12.8)
No	21 (34.0)
Missing	5 (5.3)
<b>Evaluation system uses standard, behavior-focused anchors.</b>	
On all narrative question prompts	60 (63.8)
On more than half of narrative question prompts	15 (16.0)
On half of narrative question prompts	6 (6.4)
On less than half of narrative question prompts	3 (3.2)
On none of narrative question prompts	5 (5.3)
Missing	5 (5.3)
<b>Implemented strategies to reduce bias in evaluations</b>	
Strongly agree	14 (14.9)
Agree	52 (55.3)
Neutral	15 (16.0)
Disagree	8 (8.5)
Strongly disagree	2 (2.1)
Missing	3 (3.2)
<b>Effectiveness of evaluation system at reducing bias</b>	
Very effective	12 (12.8)
Somewhat effective	47 (50.0)
Neutral	16 (17.0)
Somewhat ineffective	15 (16.0)
Very ineffective	0
Missing	4 (4.3)
<b>Reducing bias in evaluations is a priority.</b>	
Strongly agree	23 (24.5)
Agree	52 (55.3)
Neutral	13 (13.8)
Disagree	2 (2.15)
Strongly disagree	0
Missing	4 (4.3)

\*Table reports univariate statistics. Please see Results section for discussion of bivariate results.

**Table 3. Faculty Development for Core and Community Faculty About Reducing Bias in Evaluations\***

Faculty development for reducing bias	n (%)
<b>Core academic faculty preceptors attend faculty development on reducing bias in evaluations.</b>	
Twice a year	6 (6.4)
Yearly	40 (42.6)
Every other year	10 (10.6)
Less than every other year	18 (19.2)
Never	16 (17.05)
Missing	4 (4.3)
<b>Community preceptors attend faculty development on reducing bias in evaluations.</b>	
Twice a year	0
Yearly	17 (18.1)
Every other year	5 (5.3)
Less than every other year	27 (28.7)
Never	39 (41.5)
Missing	6 (6.4)
<b>Preceptors receive instructions on reducing biased language descriptors in written evaluations.</b>	
Always included in each evaluation	8 (8.5)
With each clerkship student	6 (6.4)
Once a year	30 (31.9)
Less than once a year	18 (19.2)
Never	27 (28.7)
Missing	5 (5.3)

\*Table reports univariate statistics. Please see Results section for discussion of bivariate results.

## Acknowledgments

The authors thank Lorraine Porcello for help in the literature review.

## Corresponding Author

Nina Piazza, MD

Department of Family Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY

[nina\\_piazza@urmc.rochester.edu](mailto:nina_piazza@urmc.rochester.edu)

## Author Affiliations

Nina Piazza, MD - Department of Family Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY

Mechelle Sanders, PhD - Department of Family Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY

Gerardo Moreno, MD, MS - Department of Family Medicine, University of California Los Angeles, Los Angeles, CA

Elizabeth Brown, MD, MPH - Department of Family Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY

## References

1. 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(5):684-691. doi:10.1007/s11606-019-04889-9
2. Riese A, Rappaport L, Alverson B, Park S, Rockney RM. Clinical performance evaluations of third-year medical students and association with student and evaluator gender. *Acad Med*. 2017;92(6):835-840. doi:10.1097/ACM.0000000000001565
3. 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
4. Bullock JL, Lockspeiser T, del Pino-Jones A, Richards R, Teherani A, Hauer KE. They don't see a lot of people my color: a mixed methods study of racial/ethnic stereotype threat among medical students on core clerkships. *Acad Med*. 2020;95(11S):S58-S66. doi:10.1097/ACM.0000000000003628
5. Frank AK, O'Sullivan P, Mills LM, Muller-Juge V, Hauer KE. Clerkship grading committees: the impact of group decision-making for clerkship grading. *J Gen Intern Med*. 2019;34(5):669-676. doi:10.1007/s11606-019-04879-x
6. 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
7. Chen S, Beck Dallaghan GL, Shaheen A. Implicit gender bias in third-year surgery clerkship MSPE narratives. *J Surg Educ*. 2021;78(4):1,136-1,143. doi:10.1016/j.jsurg.2020.10.011
8. Hauer KE, Lucey CR. Core clerkship grading: the illusion of objectivity. *Acad Med*. 2019;94(4):469-472. doi:10.1097/ACM.0000000000002413
9. Colson ER, Pérez M, Blaylock L, et al. Washington University School of Medicine in St. Louis case study: a process for understanding and addressing bias in clerkship grading. *Acad Med*. 2020;95(12S):S131-S135. doi:10.1097/ACM.0000000000003702
10. Onumah CM, Lai CJ, Levine D, Ismail N, Pincavage AT, Osman NY. Aiming for equity in clerkship grading: recommendations for reducing the effects of structural and individual bias. *Am J Med*. 2021;134(9):1,175-1,183.e4. doi:10.1016/j.amjmed.2021.06.001
11. Lai CJ, Alexandraki I, Ismail N, et al. Reviewing internal medicine clerkship grading through a proequity lens: results of a national survey. *Acad Med*. 2023;98(6):723-728. doi:10.1097/ACM.0000000000005142
12. Seehusen DA, Mainous AG III, Chessman AW. Creating a centralized infrastructure to facilitate medical education research. *Ann Fam Med*. 2018;16(3):257-260. doi:10.1370/afm.2228
13. Poitevien P, Osman C. Tackling implicit and explicit bias through objective structured teaching exercises for faculty. *J Grad Med Educ*. 2018;10(3):353-354. doi:10.4300/JGME-D-17-00906.1
14. Minor S, Watson KH, Bonnin R, Perkins S, Diaz V, Drowos J. Professional development needs of clerkship directors. *Fam Med*. 2023;55(5):328-332. doi:10.22454/FamMed.2023.310804
15. Society of Teachers of Family Medicine. STFM database of clerkship director characteristics. 2023.

Copyright © 2024 by the Society of Teachers of Family Medicine