




## Racial/Ethnic Disparities in Clinical Grading in Medical School

Daniel Low, Samantha W. Pollack, Zachary C. Liao, Ramoncita Maestas, Larry E. Kirven, Anne M. Eacker & Leo S. Morales


To cite this article: Daniel Low, Samantha W. Pollack, Zachary C. Liao, Ramoncita Maestas, Larry E. Kirven, Anne M. Eacker & Leo S. Morales (2019) Racial/Ethnic Disparities in Clinical Grading in Medical School, *Teaching and Learning in Medicine*, 31:5, 487-496, DOI: [10.1080/10401334.2019.1597724](https://doi.org/10.1080/10401334.2019.1597724)

To link to this article: <https://doi.org/10.1080/10401334.2019.1597724>

 [View supplementary material](#) 

 Published online: 29 Apr 2019.

 [Submit your article to this journal](#) 

 Article views: 1160


 [View related articles](#) 

 [View Crossmark data](#) 

GROUNDWORK



## Racial/Ethnic Disparities in Clinical Grading in Medical School

Daniel Low<sup>a</sup> , Samantha W. Pollack<sup>b</sup>, Zachary C. Liao<sup>c</sup>, Ramoncita Maestas<sup>d</sup>, Larry E. Kirven<sup>e</sup>, Anne M. Eacker<sup>f</sup>, and Leo S. Morales<sup>g</sup>

<sup>a</sup>Swedish Cherry Hill Family Medicine Residency, University of Washington School of Medicine, Seattle, Washington, WA, USA; <sup>b</sup>Department of Family Medicine, University of Washington School of Medicine, Seattle, Washington, USA; <sup>c</sup>Jackson Memorial Hospital, Internal Medicine Residency, University of Miami, Miami, FL, USA; <sup>d</sup>Student Affairs, University of Washington School of Medicine, Seattle, Washington, USA; <sup>e</sup>Wyoming WWAMI Program, University of Washington School of Medicine, Seattle, Washington, USA; <sup>f</sup>Kaiser Permanente School of Medicine, Pasadena, California, USA; <sup>g</sup>Center for Health Equity, Diversity, and Inclusion, University of Washington School of Medicine, Seattle, Washington, USA

### ABSTRACT

*Phenomenon:* Performance during the clinical phase of medical school is associated with membership in the Alpha Omega Alpha Honor Medical Society, competitiveness for highly selective residency specialties, and career advancement. Although race/ethnicity has been found to be associated with clinical grades during medical school, it remains unclear whether other factors such as performance on standardized tests account for racial/ethnic differences in clinical grades. Identifying the root causes of grading disparities during the clinical phase of medical school is important because of its long-term impacts on the career advancement of students of color. *Approach:* To evaluate the association between race/ethnicity and clinical grading, we examined Medical Student Performance Evaluation (MSPE) summary words (Outstanding, Excellent, Very Good, Good) and 3rd-year clerkship grades among medical students at the University of Washington School of Medicine. The analysis included data from July 2010 to June 2015. Medical students were categorized as White, underrepresented minorities (URM), and non-URM minorities. Associations between MSPE summary words and clerkship grades with race/ethnicity were assessed using ordinal logistic regression models. *Findings:* Students who identified as White or female, students who were younger in age, and students with higher United States Medical Licensing Examination Step 1 scores or final clerkship written exam scores consistently received higher final clerkship grades. Non-URM minority students were more likely than White students (Adjusted Odds Ratio = 0.53), confidence interval [0.36, 0.76],  $p = .001$ , to receive a lower category MSPE summary word in analyses adjusting for student demographics (age, gender, maternal education), year, and United States Medical Licensing Examination Step 1 scores. Similarly, in four of six required clerkships, grading disparities ( $p < .05$ ) were found to favor White students over either URM or non-URM minority students. In all analyses, after accounting for all available confounding variables, grading disparities favored White students. *Insights:* This single institution study is among the first to document racial/ethnic disparities in MSPE summary words and clerkship grades while accounting for clinical clerkship final written examinations. A national focus on grading disparities in medical school is needed to understand the scope of this problem and to identify causes and possible remedies.



### KEYWORDS


race; ethnicity; bias; grading; medical school; clerkships

## Background

Understanding predictors of clinical performance in medical school is important because clinical performance is associated with membership in the Alpha Omega Alpha (AOA) Honor Medical Society, residency selection, and career advancement.<sup>1–5</sup> Academic performance benchmarks such as preclinical grade point average<sup>6</sup> and standardized exam scores<sup>6–8</sup> have long been associated with clinical performance in

medical school. More recently, nonacademic factors, such as race/ethnicity,<sup>9</sup> clerkship order,<sup>10,11</sup> gender,<sup>12</sup> and assertiveness/extraversion,<sup>13,14</sup> have been found to be associated with clinical performance in medical school as well. However, the degree to which nonacademic factors interact with academic performance benchmarks is not well understood, as race/ethnicity has been found to be associated with medical school standardized exam scores<sup>7,15</sup> and clinical performance

**CONTACT** Daniel Low [danielhlow@gmail.com](mailto:danielhlow@gmail.com)  1959 N.E. Pacific St., Suite A-300, Seattle, WA 98195, USA. Twitter:  @DrLeoMorales; @UW; @uwmedicine

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/10401334.2019.1597724>.

independently.<sup>9</sup> There are limited data that explore the concurrent interaction of race/ethnicity with standardized exam scores and clinical performance in medical school. As a result, it remains unclear if the association between race/ethnicity and clinical performance is fully accounted for by disparities in standardized exam scores, or if there exist other, independent and potentially unrecognized factors mediating or moderating the association of race/ethnicity to clinical performance.

In a recent study of applicants to medical residencies in the United States, Black and Asian medical students were found to be less likely than their White counterparts to be members of AOA after accounting for United States Medical Licensing Examination (USMLE) Step 1 scores.<sup>16</sup> This suggests a need to further elucidate the role of race/ethnicity in the evaluation of clinical performance in medical school.

After an internal review of AOA membership at University of Washington School of Medicine (UWSOM) revealed racial/ethnic disparities, we questioned the role of race/ethnicity in the evaluation of student clerkship performance during the 3rd year because consideration for AOA membership requires that students perform in the top 25% of the class academically. We subsequently studied the association between race/ethnicity and required 3rd-year clerkship grades and Medical Student Performance Evaluation (MSPE) summary words at UWSOM.

## Methods

**Study Setting:** During required 3rd-year clerkships in Family Medicine, Internal Medicine, Obstetrics/Gynecology, Surgery, Psychiatry, and Pediatrics, students at UWSOM rotate through many clinics and hospitals at more than 100 clinical sites throughout Washington, Wyoming, Alaska, Montana, and Idaho. More than 4,000 clinical faculty, residents, and fellows evaluate students during these required 3rd-year clerkships. Demographic data for these evaluators were not available.

Third-year clerkship grades at UWSOM are based on both written final exams and clinical performance as determined by one or more clinical preceptors. Clinical performance is based on a grading rubric<sup>17</sup> with each required clinical clerkship using its own rubric specific to the skills pertinent to the particular clerkship. Upon completing a clerkship, students are given a grade of Honors, High Pass, Pass, or Fail. The weight of the written exam in final clerkship grade varied by clerkship, with some clerkships using a

threshold score for written exam to qualify for an Honors grade and others using the final written exam as a percentage of the final clerkship grade; final written exam accounted for approximately 30% to 50% of the final clerkship grade.

Upon completion of the six required 3rd-year clerkships, final grades are accumulated and weighted based on the number of Honors and High Pass grades given in each clerkship. This value is used to create a summary word that groups students into four designations: Outstanding, Excellent, Very Good, and Good. One of these four descriptors is noted in the final paragraph of each student's MSPE to summarize their overall 3rd-year performance. The MSPE is a required component of all students' residency applications.

## Participants, variables, and measures

Data for this study included MSPE summary words for all students between September 1, 2012, and June 30, 2015 ( $n = 892$ ), and the final grades of required 3rd-year clerkships (Family Medicine, Internal Medicine, Obstetrics/Gynecology, Surgery, Psychiatry, and Pediatrics) completed between July 1, 2010, and June 30, 2015 ( $n = 6,474$ ).

MSPE summary words and clerkship grades were linked to individual students. We used the American Medical College Application Service (AMCAS) application, the UWSOM Biographical and Career Preference Inventory, and the UWSOM Academic Affairs Database to abstract demographic information and academic performance data. Written clerkship exam scores were available for only four of six required clerkships, as two clerkships used different exam types across years, which prevented a comparable score from being created for analysis.

Due to the low number of clerkship Fails across all total clerkship grades ( $n = 13$ ), clerkship Pass and Fail were combined into one category, resulting in the following grade categories: Honors, High Pass, Pass/Fail. Race/ethnicity was divided into four mutually exclusive categories based on participant self-selection of race: (a) White, (b) underrepresented minority (URM; African American/Black, Latino/Hispanic of any race, American Indian or Native Alaskan, Native Hawaiians/Other Pacific Islanders), (c) non-URM minority students, and (d) missing response/declined to answer. Multiracial individuals were considered URM if any racial/ethnic category within URM was selected. Maternal education was included as a proxy measure for family socioeconomic status.<sup>18,19</sup> Given

the geographical distribution of clinical sites at UWSOM, site was divided into nine geographical categories and controlled for in statistical analyses.

### Statistical data analysis

We first evaluated the association between race/ethnicity and MSPE summary words with chi-square tests. When racial disparities were appreciated with chi-square testing, we then conducted ordinal logistic regression models to account for the ordered categories in MSPE summary words, assessing the role of race/ethnicity while accounting for age, gender, maternal education, clerkship location, clerkship year, and USMLE Step 1 in multivariate analyses. In total, we estimated 10 separate ordinal logistic regression models, one for each required 3rd-year clerkship of which six included USMLE Step 1 scores and an additional four that included clerkship final written exam scores. Only four models were estimated with written clerkship exam scores because final written clerkship exam scores were incomplete for two clerkships. USMLE Step 1 exam scores and clerkship final written exam scores were not included in the same regression models because of collinearity. Maternal education, geographical distribution of clinical sites, gender, and age were included as covariates due to a priori knowledge and literature review demonstrating they can affect medical school performance. Father's education was excluded due to collinearity with mother's education. Because we wanted to assess antecedent factors impacting clerkship performance, we did not include Step 2 examinations.

Clerkships have been randomly numbered 1 to 6 in this analysis. All calculations were done using SPSS Version 18 (SPSS Inc., Chicago). The University of Washington Institutional Review Board approved this research under FWA #00006878 (IRB ID: STUDY00001562).

## Results

### Descriptive analyses

There were 1,096 students who received 6,474 assessments in Family Medicine, Internal Medicine, Psychiatry, Obstetrics/Gynecology, Pediatrics, and Surgery (Table 1). Not all students completed all six required clerkships over the study period, resulting in an unequal number of grades per required clerkship. Among the 1,096 students in the study, 66% were White, 8% were URM, 15% were non-URM minorities, and 11% had missing race/ethnicity information.

**Table 1.** Characteristics of study participants at the University of Washington School of Medicine

Age Quartiles <i>n</i> (%)	
≤ 24 years	136 (12)
25–28 years	677 (62)
29–32 years	209 (19)
≥ 33 years	74 (7)
Gender <i>n</i> (%)	
Male	489 (45)
Female	607 (55)
Race/Ethnicity <i>n</i> (%)	
White	720 (66)
URM <sup>1</sup>	87 (8)
Non-URM Minority	167 (15)
Unknown/Missing	122 (11)
Maternal Education <i>n</i> (%)	
≤ High School	143 (13)
Post High School/Community College	185 (17)
Bachelor's Degree	287 (26)
BA+/Master's Degree	313 (29)
PhD or Above	117 (11)
Missing	51 (5)
Clerkship Year <i>n</i> (%)	
2010–2011	227 (21)
2011–2012	220 (20)
2012–2013	224 (20)
2013–2014	214 (20)
2014–2015	211 (19)
USMLE Step 1 <i>M</i> ( <i>SD</i> ) [Range]	226 (19) [159–274]
Clerkship Exam Mean Scores	
Clerkship 1	79 (7.7) [59–99]
Clerkship 2	85 (5.8) [66–98]
Clerkship 3	82 (7.4) [44–100]
Clerkship 4	76 (7.4) [56–99]

Note: Underrepresented minority (URM) includes African American/Black, Hispanic/Latino, American Indian/Alaskan Native and Native Hawaiian/Other Pacific Islander students. USMLE = US Medical Licensing Examination.

The average age was 27 years (range = 21–48), and 55% were female.

### MSPE summary words and clerkship grades

Table 2 shows the distribution of MSPE summary words. Overall, 27% of students received Outstanding, 31% received Excellent, 31% received Very Good, and 11% received Good. The highest proportion of students receiving Outstanding were White students (71%), followed by non-URM minority students (13%). Only 3% of all Outstanding MSPE summary words were earned by URM students,  $\chi^2 = 29.92$ ,  $p < .001$ . Table 3 shows the distributions of grades by race/ethnicity and clerkship. White students received the highest percentage of Honors across all clerkships (34%–46%), followed by non-URM minority students (29%–39%) and URM students (16%–40%).

### Ordinal logistic regression results

#### MSPE summary word assignments

Results from the analyses of the MSPE summary word assignments are shown in Figure 1. In multivariate

**Table 2.** Distribution of Medical Student Performance Evaluation summary word by race/ethnicity at University of Washington School of Medicine

	Outstanding <i>n</i> (%)	Excellent <i>n</i> (%)	Very Good <i>n</i> (%)	Good <i>n</i> (%)	Total
Students, <i>n</i>	243	272	280	97	892
Race/Ethnicity					
White	173 (71)	181 (67)	178 (64)	48 (49)	580 (65)
URM <sup>1</sup>	7 (3)	17 (6)	28 (10)	14 (14)	66 (7)
Non-URM Minority	31 (13)	45 (17)	42 (15)	25 (26)	143 (16)
Missing/Declined	32 (13)	29 (11)	32 (11)	10 (10)	103 (12)

Note:  $\chi^2(9, N = 892) = 29.92, p < .001$ . Underrepresented minority (URM) includes African American/Black, Hispanic/Latino, American Indian/Alaskan Native and Native Hawaiian/Other Pacific Islander students.

**Table 3.** Clinical grades by race/ethnicity and clerkships at University of Washington School of Medicine

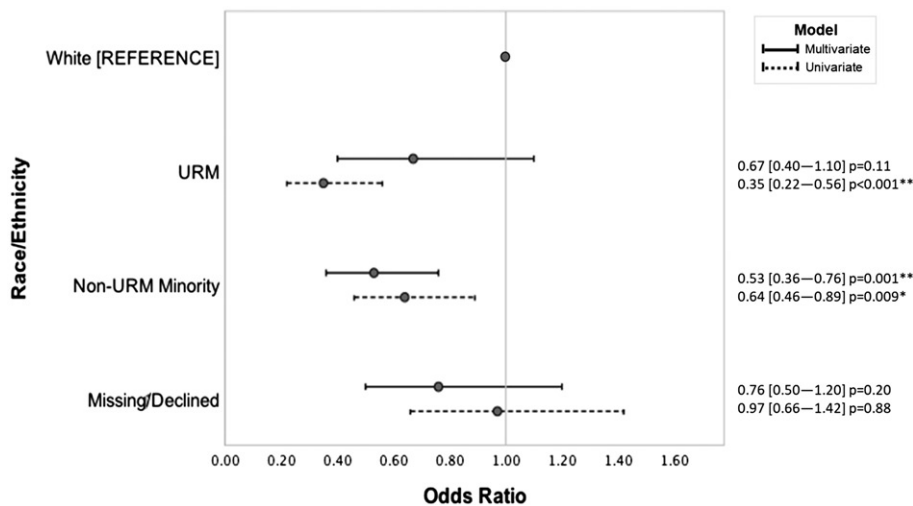
	Required Third-Year Clerkships					
	Clerkship 1 <sup>a</sup> <i>n</i> (%)	Clerkship 2 <sup>a</sup> <i>n</i> (%)	Clerkship 3 <sup>b</sup> <i>n</i> (%)	Clerkship 4 <sup>a</sup> <i>n</i> (%)	Clerkship 5 <sup>c</sup> <i>n</i> (%)	Clerkship 6 <sup>a</sup> <i>n</i> (%)
White						
Honors	271 (38)	284 (40)	242 (34)	267 (38)	326 (46)	271 (38)
High Pass	300 (42)	351 (49)	355 (50)	265 (37)	344 (49)	216 (31)
Pass/Fail	139 (20)	75 (11)	109 (15)	179 (25)	39 (6)	222 (31)
URM						
Honors	14 (16)	24 (28)	17 (20)	17 (20)	34 (40)	23 (27)
High Pass	35 (41)	42 (49)	43 (51)	28 (33)	44 (52)	24 (28)
Pass/Fail	37 (43)	19 (22)	25 (29)	39 (46)	7 (8)	38 (45)
Non-URM Minority						
Honors	52 (32)	51 (31)	48 (29)	47 (29)	65 (39)	56 (34)
High Pass	75 (46)	87 (52)	80 (49)	53 (32)	82 (50)	50 (30)
Pass/Fail	37 (23)	28 (17)	36 (22)	65 (39)	18 (11)	60 (36)
Missing/Declined						
Honors	53 (44)	45 (38)	36 (30)	53 (44)	58 (48)	47 (39)
High Pass	49 (41)	60 (50)	67 (56)	35 (29)	56 (47)	23 (19)
Pass/Fail	18 (15)	14 (12)	17 (14)	32 (27)	6 (5)	50 (42)

Note: Underrepresented minority (URM) includes African American/Black, Hispanic/Latino, American Indian/Alaskan Native and Native Hawaiian/Other Pacific Islander students.

<sup>a</sup>*N* = 1,080.

<sup>b</sup>*N* = 1,075.

<sup>c</sup>*N* = 1,079.



**Figure 1.** Univariate and multivariate adjusted odds ratios and 95% CIs for the association of MSPE summary word and race/ethnicity. Note. Multivariate model includes gender, age, maternal education, clerkship year, and USMLE Step 1 score. \**p* < .05. \*\**p* < .001.

analysis, non-URM minority students were significantly less likely to receive a higher category word than White students (e.g., Outstanding vs. Excellent; adjusted odds ratio [AOR]=0.53), 95% confidence interval (CI) [0.36, 0.76], *p* = .001, whereas URM

status trended toward being less likely to receive a higher category word, but without statistical significance (AOR = 0.67), 95% CI [0.40, 1.10], *p* = .11.

Men were less likely to receive a higher category MSPE summary word than women (AOR = 0.46), CI

**Table 4.** Odds of receiving a higher Medical Student Performance Evaluation summary word at University of Washington School of Medicine

	Univariate Model		Multivariate Model <sup>a</sup>	
	OR [95% CI]	<i>p</i>	AOR [95% CI]	<i>p</i>
Race		<.001		.01
[White = REF]	1.00		1.00	
URM	0.35 [0.22, 0.56]	<.001	0.67 [0.40, 1.10]	.11
Non-URM Minority	0.64 [0.46, 0.89]	.009	0.53 [0.36, 0.76]	.001
Missing/Declined	0.97 [0.66, 1.42]	.88	0.76 [0.50, 1.16]	.20
Gender				
[Female = REF]	1.00		1.00	
Male	0.75 [0.59, 0.95]	.02	0.46 [0.35, 0.60]	<.001
Age	0.90 [0.86, 0.93]	<.001	0.93 [0.89, 0.97]	<.001
Step1	1.07 [1.06, 1.07]	<.001	1.08 [1.07, 1.08]	<.001
Maternal Education		<.001		.33
≤ High School [REF]	1.00		1.00	
Post High School/Community College	1.58 [1.00, 2.48]	.045	1.05 [0.65, 1.70]	.85
Bachelor's Degree	2.31 [1.51, 3.53]	<.001	1.23 [0.78, 1.94]	.38
BA+/Master's Degree	1.90 [1.25, 2.88]	.003	1.39 [0.89, 2.19]	.15
Ph.D. or Above	2.94 [1.78, 4.87]	<.001	1.55 [0.90, 2.66]	.11
Clerkship Year <i>n</i> (%)		.15		<.001
2014–2015 [REF]	1.00		1.00	
2010–2011	0.75 [0.37, 1.49]	.41	2.29 [1.09, 4.82]	.03
2011–2012	1.41 [1.00, 1.99]	.05	3.37 [2.26, 5.03]	<.001
2012–2013	1.03 [0.73, 1.44]	.88	1.70 [1.16, 2.50]	.01
2013–2014	1.23 [0.87, 1.74]	.25	1.66 [1.13, 2.43]	.01

Note: OR = odds ratio; CI = confidence interval; AOR = adjusted OR; REF = XXXX; URM = underrepresented minority.

<sup>a</sup>Includes gender, age, maternal education, clerkship year, and Step 1 score.

[0.35, 0.60],  $p < .001$ , as were older students (AOR = 0.93), CI [0.89, 0.97],  $p < .001$ . Higher USMLE Step 1 exam scores were associated with higher category MSPE summary word (AOR = 1.08), CI [1.07, 1.08],  $p < .001$  (Table 4).

### Required clerkship grades

Table 5 shows results from the ordinal logistic regression models of clerkship grades. In multivariate models including USMLE Step 1 scores, URM students were less likely than their White counterparts to receive a higher clerkship grade in one of six clerkships—Clerkship 1 (AOR = 0.49), 95% CI [0.30, 0.78],  $p = .003$ —and non-URM minority students were less likely to receive a higher clerkship grade in four of six clerkships: Clerkship 1 (AOR = 0.69), 95% CI [0.48, 0.99],  $p = .05$ ; Clerkship 2 (AOR = 0.56), 95% CI [0.40, 0.80],  $p = .001$ ; Clerkship 3 (AOR = 0.69), 95% CI [0.48, 0.99],  $p = .04$ ; and Clerkship 4 (AOR = 0.58), 95% CI [0.41, 0.82],  $p = .002$ . In multivariate models including a written final exam score (but not USMLE Step 1 scores), URM students were less likely than their White counterparts to receive a higher clerkship grade in one of four clerkships—Clerkship 1 (AOR = 0.59), 95% CI [0.35, 0.96],  $p = .03$ —and non-URM minority students were less likely to receive a higher clerkship grade in one of four clerkships—Clerkship 2 (AOR = 0.62), 95% CI [0.43, 0.90],  $p = .01$ . Students missing racial/ethnic data did not differ from white students in the results of any model.

Age, gender, clerkship year, clerkship location, USMLE Step 1 score, and final clerkship exam scores were all associated with final clerkship grades in at least one of six required clerkships in multivariate models. There were no consistent associations with clerkship grades and clerkship year or location; however, age, gender, USMLE Step 1 score, and final clerkship exam scores were consistently associated with clerkship grades. Female participants, younger students, and those with higher USMLE Step 1 scores and final clerkship exam scores consistently received higher final clerkship grades. Maternal education had no association with final clerkship grades in any clerkship.

### Discussion

Required 3rd-year clerkship grades are recognized as one of the most important factors residency programs consider in selecting residents.<sup>1</sup> In this study, we examined the association between race/ethnicity and clinical clerkship grades at UWSOM. Using multivariate ordinal logistic regression analyses, we found that race/ethnicity, gender, age, clerkship final written examination scores, and USMLE Step 1 exam scores were repeatedly and independently associated with clinical clerkship grades for 3rd-year required clerkships.

As reported in prior research, higher USMLE Step 1 exam scores<sup>6</sup> and female gender<sup>12</sup> were associated

**Table 5.** Odds of receiving a higher clerkship grade at University of Washington School of Medicine

	Univariate Model		Adjusted for All Covariates <sup>a</sup> + USMLE Step 1		Adjusted for All Covariates <sup>a,b</sup> + Final Exam Score	
	OR [95% CI]	<i>p</i>	AOR [95% CI]	<i>p</i>	AOR [95% CI]	<i>p</i>
<b>Clerkship 1<sup>c</sup></b>						
Race		<.001		.003		.05
White [REF]	1.00		1.00		1.00	
URM	0.32 [0.21, 0.49]	<.001	0.49 [0.30, 0.78]	.003	0.59 [0.35, 0.96]	.03
Non-URM Minority	0.78 [0.57, 1.08]	.13	0.69 [0.48, 1.00]	.05	0.85 [0.58, 1.26]	.43
Missing/Declined	1.31 [0.91, 1.88]	.15	1.24 [0.82, 1.87]	.30	1.40 [0.89, 2.19]	.15
USMLE Step 1	1.07 [1.06, 1.08]	<.001	1.07 [1.06, 1.08]	<.001	NA	
Final Exam Score	1.30 [1.27, 1.33]	<.001	NA		1.31 [1.27, 1.35]	<.001
Gender						
Female [REF]	1.00		1.00		1.00	
Male	0.90 [0.72, 1.13]	.36	0.58 [0.45, 0.75]	<.001	0.69 [0.52, 0.90]	.007
Age	0.89 [0.86, 0.92]	<.001	0.92 [0.88, 0.95]	<.001	0.94 [0.90, 0.98]	.003
<b>Clerkship 2<sup>c</sup></b>						
Race		.003		.01		.09
White [REF]	1.00		1.00		1.00	
URM	0.51 [0.33, 0.79]	.003	0.72 [0.45, 1.16]	.17	0.97 [0.60, 1.58]	.91
Non-URM Minority	0.64 [0.46, 0.89]	.007	0.56 [0.40, 0.80]	.001	0.62 [0.43, 0.90]	.01
Missing/Declined	0.91 [0.63, 1.32]	.61	0.81 [0.55, 1.21]	.30	0.92 [0.61, 1.40]	.70
USMLE Step 1	1.03 [1.02, 1.04]	<.001	1.03 [1.03, 1.04]	<.001	NA	
Final Exam Score	1.21 [1.18, 1.24]	<.001	NA		1.22 [1.19, 1.25]	<.001
Gender						
Female [REF]	1.00		1.00		1.00	
Male	0.69 [0.54, 0.86]	.001	0.53 [0.41, 0.68]	<.001	0.79 [0.61, 1.02]	.07
Age	0.93 [0.89, 0.96]	<.001	0.94 [0.91, 0.98]	.002	0.94 [0.90, 0.98]	.001
<b>Clerkship 3<sup>d</sup></b>						
Race		.002		.15		.68
White [REF]	1.00		1.00		1.00	
URM	0.46 [0.30, 0.70]	<.001	0.71 [0.45, 1.13]	.15	0.91 [0.52, 1.58]	.73
Non-URM Minority	0.73 [0.53, 1.01]	.06	0.69 [0.48, 0.99]	.04	1.05 [0.68, 1.61]	.84
Missing/Declined	0.91 [0.63, 1.30]	.60	0.91 [0.61, 1.35]	.63	0.76 [0.47, 1.23]	.26
USMLE Step 1	1.04 [1.03, 1.05]	<.001	1.05 [1.04, 1.05]	<.001	NA	
Final Exam Score	1.43 [1.38, 1.47]	<.001	NA		1.46 [1.41, 1.52]	<.001
Gender						
Female [REF]	1.00		1.00		1.00	
Male	0.86 [0.68, 1.08]	.19	0.60 [0.47, 0.78]	<.001	0.94 [0.70, 1.26]	.67
Age	0.94 [0.91, 0.97]	.001	0.98 [0.95, 1.02]	.36	0.99 [0.94, 1.04]	.66
<b>Clerkship 4<sup>c</sup></b>						
Race		<.001		.01		.33
White [REF]	1.00		1.00		1.00	
URM	0.41 [0.27, 0.62]	<.001	0.68 [0.43, 1.07]	.10	0.85 [0.52, 1.39]	.85
Non-URM Minority	0.58 [0.42, 0.80]	.001	0.58 [0.41, 0.82]	.002	0.81 [0.56, 1.17]	.25
Missing/Declined	1.15 [0.80, 1.66]	.45	1.07 [0.72, 1.59]	.73	1.28 [0.84, 1.94]	.26
USMLE Step 1	1.04 [1.03, 1.05]	<.001	1.04 [1.04, 1.05]	<.001	NA	
Final Exam Score	1.23 [1.20, 1.26]	<.001	NA		1.23 [1.20, 1.26]	<.001
Gender						
Female [REF]	1.00		1.00		1.00	
Male	0.38 [0.54, 0.84]	.001	0.51 [0.40, 0.65]	<.001	0.77 [0.60, 1.00]	.05
Age	0.96 [0.93, 1.00]	.03	1.01 [0.97, 1.05]	.54	1.04 [1.00, 1.08]	.07
<b>Clerkship 5<sup>e</sup></b>						
Race		.11		.34		
White [REF]	1.00		1.00		1.00	
URM	0.76 [0.49, 1.19]	.23	0.94 [0.58, 1.53]	.81		
Non-URM Minority	0.70 [0.50, 0.98]	.04	0.71 [0.49, 1.03]	.07		
Missing/Declined	1.10 [0.75, 1.60]	.63	0.89 [0.58, 1.36]	.59		
USMLE Step 1	1.03 [1.02, 1.03]	<.001	1.03 [1.02, 1.04]	<.001		
Gender						
Female [REF]	1.00		1.00		1.00	
Male	0.59 [0.47, 0.75]	<.001	0.49 [0.38, 0.64]	<.001		
Age	0.94 [0.91, 0.98]	.001	0.98 [0.94, 1.02]	.34		
<b>Clerkship 6<sup>c</sup></b>						
Race		.05		.35		
White [REF]	1.00		1.00		1.00	
URM	0.58 [0.38, 0.89]	.01	0.90 [0.56, 1.43]	.65		
Non-URM Minority	0.82 [0.60, 1.11]	.20	0.92 [0.65, 1.30]	.63		
Missing/Declined	0.81 [0.56, 1.17]	.27	0.69 [0.47, 1.04]	.08		
USMLE Step 1	1.03 [1.03, 1.04]	<.001	1.04 [1.03, 1.05]	<.001		
Gender						
Female [REF]	1.00		1.00		1.00	
Male	0.84 [0.67, 1.04]	.11	0.67 [0.52, 0.85]	.001		
Age	0.92 [0.88, 0.95]	<.001	0.95 [0.91, 0.98]	.01		

Note: USMLE = United States Medical Licensing Examination; OR = odds ratio; CI = confidence interval; AOR = adjusted OR; REF = XXXX; URM = underrepresented minority.

<sup>a</sup>Includes gender, age, clerkship year, maternal education, and clerkship region.

<sup>b</sup>No final exam score available for Clerkships 5 and 6, so they have only one multivariate model.

<sup>c</sup>*N* = 1,080.

<sup>d</sup>*N* = 1,075.

<sup>e</sup>*N* = 1,079.

with higher clinical clerkship grades. We anticipated higher exam scores correlating with higher grades and MSPE summary words, as the exams are proxies for clinical knowledge, and thus it is to be expected that those with greater knowledge would do better clinically. We similarly anticipated female gender to be associated with higher marks, as female gender has repeatedly been found in medicine to be associated with higher degrees of empathy and interpersonal skills, better clinical performance,<sup>20–22</sup> and better medical school grades.<sup>12</sup> The difference observed in age is not easily explainable; other U.S. data have shown no significant difference in clinical performance by age.<sup>23</sup> There are many possible reasons that older students in our cohort were less likely to receive higher marks, meriting further research. Possible reasons could include older students having more responsibilities outside the classroom, older students having had greater time away from an academic setting, or older students having different interpersonal dynamics with evaluators.

We also found that all non-White students (both URM and non-URM) received lower final clerkship grades than White students even after adjusting for the aforementioned factors. Non-URM minority students also received lower MSPE summary words than White students. To our knowledge, this is the first report of these findings within a U.S. medical school that includes race/ethnicity, USMLE Step 1 exam scores, and final written clinical clerkship exam scores in the analyses of clerkship grades.

Although some of the race/ethnicity differences in clerkship assessments disappeared after accounting for test scores—indicating the importance of medical knowledge tested by such examinations—racial/ethnic disparities nonetheless persisted in four of six required clerkships after accounting for many possible confounders. Observing such racial/ethnic disparities in grades after adjusting for exam scores and other factors suggests that grade disparities may be attributable to differences in grading independent of student clinical performance. Specifically, instructor bias may be contributing to the observed disparities in grades. A substantial body of research demonstrate that physicians hold implicit and explicit biases favoring Whites over racial/ethnic minorities.<sup>24–28</sup> This is important in understanding our results, given all observed racial/ethnic disparities were unidirectional, with URM or non-URM minority students receiving lower grades than White students, whereas White students never performed lower than students of color. The observed grading disparities therefore might be a manifestation of previously

reported racial/ethnic biases. This aligns with the qualitative experience of medical students of color, who report increased discrimination compared to Whites.<sup>29,30</sup> Such an explanation would also align with previously observed racial/ethnic disparities in academic medicine, where minority medical faculty have been found less likely to be promoted than Whites after accounting for peer-reviewed publications, research funding, clinical activity, and tenure status,<sup>31–33</sup> and Black physician-scientists have been found less likely to receive National Institutes of Health funding than Whites after accounting for relevant qualifications.<sup>34,35</sup>

Other explanations for the observed grading disparities are also possible. External factors such as students' personal circumstances, including family demands and financial stressors, may contribute to clerkship grades.<sup>36</sup> A racialized learning environment might prevent URM and non-URM minority students from performing their best.<sup>30,37</sup> Instructor characteristics such as race/ethnicity, gender, and age may also contribute to disparities. Individual student–teacher interactions may be an important variable and warrants further study, for despite the use of grading rubrics, the evaluation of clinical performance at most medical schools, including UWSOM, is largely based on instructors' subjective observations of students over the course of a clerkship.

The observed disparities are alarming because the receipt of clerkship grades, MSPE summary words, and AOA membership influences the career trajectory of medical students.<sup>1–5</sup> As recent evidence shows, even marginal differences in assessed clinical performance can lead to markedly increased consequences for students underrepresented in medicine.<sup>38</sup> Our results call into question the objectivity and validity of what is measured by clinical performance evaluations. Instructor biases may be influencing clinical grading and contributing to a racialized learning environment that harms students of color, suggesting a need to develop approaches that reduce the influence of implicit and explicit bias on clinical evaluations.

It therefore is imperative for all medical schools to critically examine their current evaluation practices. To this end, several best practices are available, including the use of evaluation rubrics and checklists to reduce subjectivity,<sup>39</sup> a slowdown in the decision process to encourage analytic thinking, education of faculty on the existence of implicit bias and its possible unintentional and imperceptible influence on assessment, and other practices that aim for objective assessment.<sup>40</sup> Research shows that African American physicians hold less implicit race bias than others,<sup>24</sup> which suggest that



efforts to increase faculty racial diversity may also improve equity in clerkship grading. Competency-based assessments with the elimination of distinction may also minimize the impact of implicit and explicit biases by instructors. In addition, developing meaningful policies within academia that reward assessment of clinical performance by faculty (e.g., precepting could count toward productivity and/or be included as criteria for promotion) may augment the systematic training of faculty in evaluating medical students.<sup>41,42</sup>

At UWSOM we have begun this process with improved standardization of the grading process across clinical clerkships, implicit bias education for clinical faculty, and antiracism training for all senior medical school administrators and residency program directors. We also have formed an antiracism action committee to address learning environment concerns, developed an active program to increase faculty diversity, and continued our ongoing evaluation of clinical assessments. It is also important for residency programs to consider the impact of biases present in medical schools as in our society when making decisions about ranking residents.

This research has limitations that need be considered when interpreting its results. This is a study of a single institution; results may not generalize to other medical schools. Our limited number of URM students may underestimate disparities due to a lack of power and explain the wide confidence intervals in many of our analyses. Due to low numbers of many specific racial/ethnic identities, and the limitation of racial/ethnic categories offered by AMCAS during the study period, we combined very diverse racial/ethnic identities into URM and non-URM categories, even though the heterogeneity of individuals composing these groups cannot appropriately be assigned by such categorization. Missing data for race/ethnicity in 11% of participants limit full assessment. The data are observational and limited to the available variables; notably, assessment of clinical performance was not available. Discordant testing results from some clerkships limited uniform data collection. Clinical sites were grouped in analysis, so variance by specific site was not examined. Other, unknown student and faculty factors may contribute to grading disparities. The data included in this study predate the new curriculum at UWSOM, which was implemented in 2015 for entering 1st-year students.

## Conclusion

We found disparities in the required 3rd-year clerkship grades and in the MSPE summary words at

UWSOM. Some of these disparities were anticipated; higher national exam scores are associated with higher grades. Other differences, such as female gender, have been reported previously. This research highlights the racial/ethnic disparities in clinical grades after accounting for the preceding demographic and test score differences. Racial/ethnic bias may contribute to the grading disparities observed in clinical performance, necessitating a comprehensive review of how assessments of clinical performance in medical school are derived and suggesting residency programs review clerkship grades of URM and non-URM minority students with caution, due to unaccounted for disparities that may be secondary to evaluator bias. A focus on grading disparities in medical school is needed to understand the scope of this problem nationally and to identify causes and possible remedies.

## Acknowledgments

We thank the medical students of the University of Washington School of Medicine for contributing their data to this study and thank the medical school's administrators, specifically, Suzanne Allen, M.D., Vice Dean for Academic, Rural and Regional Affairs, and Paul Ramsey, M.D., Dean of the UW School of Medicine, for providing access to the school's data systems for this study. We also thank Janice Sabin, Ph.D., Research Associate Professor, for her comments on the manuscript, and the Department of Biomedical Informatics and Medical Education for supporting Samantha Pollack's time on this project. Finally, we thank MaiLinh Niemi for her help in preparing this manuscript.

## ORCID

Daniel Low  <http://orcid.org/0000-0003-0555-2376>

## References

1. Green M, Jones P, Thomas JX. Jr., Selection criteria for residency: results of a national program directors survey. *Acad Med.* 2009;83(3):362–367. doi:10.1097/ACM.0b013e3181970c6b.
2. LaGrasso JR, Kennedy DA, Hoehn JG, et al. Selection criteria for the integrated model of plastic surgery residency. *Plast Reconstr Surg.* 2008;121(3):121e–125e. doi:10.1097/01.prs.0000299456.96822.1b.
3. Rinard JR, Garol BD, Shenoy AB, Mahabir RC. Successfully matching into surgical specialties: an analysis of national resident matching program data. *J Grad Med Educ.* 2010;2(3):316–321. doi:10.4300/JGME-D-09-00020.1.
4. DeCroff CM, Mahabir RC, Zamboni WA. The impact of alpha omega alpha membership on successfully matching to residency. *Plast Reconstr Surg.* 2010;126(2):113e–115e. doi:10.1097/PRS.0b013e3181df70b3.

5. National Resident Matching Program. Charting outcomes in the match. Characteristics of applicants who matched to their preferred specialty in the 2014 main residency match. 2014. <http://www.nrmp.org/wp-content/uploads/2014/09/Charting-Outcomes-2014-Final.pdf>. Accessed March 24, 2017.
6. Casey PM, Palmer BA, Thompson GB. Predictors of medical school clerkship performance: a multispecialty longitudinal analysis of standardized examination scores and clinical assessments. *BMC Med Educ* 2016; 16:128. doi:10.1186/s12909-016-0652-y.
7. Donnon T, Paolucci EO, Violato C. The predictive validity of the MCAT for medical school performance and medical board licensing examinations: a meta-analysis of the published research. *Acad Med*. 2007; 82(1):100–106. doi:10.1097/01.ACM.0000249878.25186.b7.
8. Ferguson E, James D, Madeley L. Factors associated with success in medical school: systematic review of the literature. *BMJ*. 2002; 324(7343):952–957. doi:10.1136/bmj.324.7343.952.
9. Lee KB, Vaishnavi SN, Lau SK, et al. Making the grade:” noncognitive predictors of medical students’ clinical clerkship grades. *J Nat Med Assoc* 2007;99(10): 1138–1150.
10. Hampton HL, Collins BJ, Perry KG, Jr, et al. Order of rotation in third-year clerkships. Influence on academic performance. *J Reprod Med*. 1996;41(5): 337–340.
11. Reteguiz JA, Crosson J. Clerkship order and performance on family medicine and internal medicine National Board of Medical Examiners Exams. *Fam Med*. 2002;34(8):604–608.
12. Riese A, Rappaport L, Alverson B, et al. 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.
13. Shen H, Comrey AL. Factorial validity of personality structure in medical school applicants. *Educ Psychol Meas*. 1995;55(6):1008–1015. doi:10.1177/0013164495055006010.
14. Davis KR, Banken JA. Personality type and clinical evaluations in an obstetrics/gynecology medical student clerkship. *Am J Obstet Gynecol*. 2005;193(5): 1807–1810. doi:10.1016/j.ajog.2005.07.082.
15. Davis D, Dorsey JK, Franks RD, et al. Do racial and ethnic group differences in performance on the MCAT exam reflect test bias?. *Acad Med*. 2013;88(5): 593–602. doi:10.1097/ACM.0b013e318286803a.
16. Boatright D, Ross D, O’Connor P, et al. Racial disparities in medical student membership in the alpha omega alpha honor society. *JAMA Intern Med*. 2017; 177(5):659–665. doi:10.1001/jamainternmed.2016.9623.
17. Pangaro LN. A new vocabulary and other innovations for improving descriptive in-training evaluations. *Acad Med*. 1999;74(11):1203–1207.
18. Adler NE, Boyce T, Chesney MA, et al. Socioeconomic status and health. The challenge of the gradient. *Am Psychol*. 1994;49(1):15–24. Review.
19. Psaki SR, Seidman JC, Miller M, et al. Measuring socioeconomic status in multicountry studies: results from the eight-country MAL-ED study. *Popul Health Metr* 2014;12(1):8. doi:10.1186/1478-7954-12-8.
20. Swygert KA, Cuddy MM, van Zanten M, et al. Gender differences in examinee performance on the step 2 clinical skills data gathering (DG) and patient note (PN) components. *Adv in Health Sci Educ*. 2012; 17(4):557–571. doi:10.1007/s10459-011-9333-0.
21. Austin EJ, Evans P, Goldwater R, et al. A preliminary study of emotional intelligence, empathy and exam performance in first year medical students. *Pers Individ Dif*. 2005;39(8):1395–1405. doi:10.1016/j.paid.2005.04.014.
22. Hojat M, Gonnella JS, Mangione S, et al. Empathy in medical students as related to academic performance, clinical competence and gender. *Med Educ*. 2002; 36(6):522–527. doi:10.1046/j.1365-2923.2002.01234.x.
23. Cortez AR, Winer LK, Kim Y, et al. Predictors of medical student success on the surgery clerkship. *Am J Surg*. 2019; 217(1):169–174. doi:10.1016/j.amjsurg.2018.09.021.
24. Sabin JA, Nosek BA, Greenwald AG, et al. Physicians’ implicit and explicit attitudes about race by MD race, ethnicity, and gender. *J Health Care Poor Underserved*. 2009;20(3):896–913. doi:10.1353/hpu.0.0185.
25. Green AR, Carney DR, Pallin DJ, et al. Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. *J Gen Intern Med*. 2007;22(9):1231–1238. doi:10.1007/s11606-007-0258-5.
26. Penner LA, Dovidio JF, West TV, et al. Aversive racism and medical interactions with black patients: a field study. *J of Experimental Social Psychology*. 2010; 46(2):436–440. doi:10.1016/j.jesp.2009.11.004.
27. Blair IV, Steiner JF, Hanratty R, et al. An investigation of associations between clinicians’ ethnic or racial bias and hypertension treatment, medication adherence and blood pressure control. *J Gen Intern Med*. 2014;29(7):987–995. doi:10.1007/s11606-014-2795-z.
28. Haider AH, Sexton J, Sriram N, et al. Association of unconscious race and social class bias with vignette-based clinical assessments by medical students. *JAMA* 2011;306(9):942–951.
29. Kassebaum DG, Cutler ER. On the culture of student abuse in medical school. *Acad Med*. 1998;73(11): 1149–1158.
30. Orom H, Semalulu T, Underwood W. The social and learning environments experienced by underrepresented minority medical students: a narrative review. *Acad Med*. 2013;88(11):1765–1777. doi:10.1097/ACM.0b013e3182a7a3af.
31. Nunez-Smith M, Ciarleglio MM, Sandoval-Schaefer T, et al. Institutional variation in the promotion of racial/ethnic minority faculty at US medical schools. *Am J Public Health*. 2012;102(5):852–858. doi:10.2105/AJPH.2011.300552.
32. Fang D, Moy E, Colburn L, Hurley J. Racial and ethnic disparities in faculty promotion in academic medicine. *JAMA* 2000;284(9):1085–1092.
33. Palepu A, Carr PL, Friedman RH, et al. Minority faculty and academic rank in medicine. *JAMA* 1998; 280(9):767–771.

34. Ginther DK, Haak LL, Schaffer WT, et al. Are race, ethnicity, and medical school affiliation associated with NIH R01 type 1 award probability for physician investigators?. *Acad Med.* 2012;87(11):1516–1524. doi: [10.1097/ACM.0b013e31826d726b](https://doi.org/10.1097/ACM.0b013e31826d726b).
35. Ginther DK, Schaffer WT, Schnell J, et al. Race, ethnicity, and NIH research awards. *Science.* 2011;333(6045):1015–1019. doi:[10.1126/science.1196783](https://doi.org/10.1126/science.1196783).
36. Abdulghani HM, Al-Drees AA, Khalil MS, et al. What factors determine academic achievement in high achieving undergraduate medical students? A qualitative study. *Med Teach.* 2014;36 Suppl 1:S43–S48. doi: [10.3109/0142159X.2014.886011](https://doi.org/10.3109/0142159X.2014.886011).
37. Steele CM, Aronson J. Stereotype threat and the intellectual test performance of African Americans. *J Pers Soc Psychol.* 1995;69(5):797–811.
38. Teherani A, Hauer KE, Fernandez A, et al. How small differences in assessed clinical performance amplify to large differences in grades and awards: a cascade with serious consequences for students underrepresented in medicine. *Acad Med.* 2018;93(9):1286–1292. doi: [10.1097/ACM.0000000000002323](https://doi.org/10.1097/ACM.0000000000002323).
39. Ely JW, Graber ML, Croskerry P. Checklists to reduce diagnostic errors. *Acad Med.* 2011;86(3):307–313. doi: [10.1097/ACM.0b013e31820824cd](https://doi.org/10.1097/ACM.0b013e31820824cd).
40. Croskerry P, Singhal G, Mamede S. Cognitive debiasing 2: impediments to and strategies for change. *BMJ Qual Saf.* 2013;22(Suppl 2):ii65–ii72. doi:[10.1136/bmjqs-2012-001713](https://doi.org/10.1136/bmjqs-2012-001713).
41. Holmboe ES, Hawkins RE, Huot SJ. Effects of training in direct observation of medical residents' clinical competence: a randomized trial. *Ann Intern Med.* 2004;140(11):874–881.
42. Holmboe ES. Faculty and the observation of trainees' clinical skills: problems and opportunities. *Acad Med.* 2004;79(1):16–22.