Supplementary Online Content

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods

Application Review Process

Department staff manually screened each application, using Acrobat (Adobe, San Jose, California) to redact applicant identifiers that might trigger implicit bias—name, sex/gender (including pronouns signifying gender); race/ethnicity; and affiliation with race-, gender-, or religion-associated groups or activities (e.g., Latino Medical Student Association). Descriptions of leadership positions, roles, and activities within redacted associations were not altered. Redaction encompassed full applications, including essays and accompanying letters. The residency program director and coordinator reviewed a random sample of redacted applications to ensure accuracy. Each application was reviewed in redacted form by one reviewer and unredacted by a second reviewer.

All department faculty participate in screening, with each application independently reviewed by two randomly-selected faculty members who have received implicit bias training. Faculty reviewers each assign a score, all applications and preliminary scores are reviewed by the program director and department chair, and interview invitations extended accordingly.

Forty-six faculty reviewed 462 applications (19-23 applications each). Reviewer assignments were randomized to minimize confounding from within-reviewer effects; no pair of reviewers reviewed the same set of applications except as occurred by chance. Each faculty member reviewed a mix of redacted and unredacted applications. Applications were scored from 1 (best) to 9 (worst), relative to the usual spectrum of applications. Reviewers were instructed that most scores should fall mid-range, with 5 representing our average applicant (strong overall), and that it would be unusual to for each reviewer to score more than one to two of their respective set of applications as "1." Faculty were given examples to aid review:

> 1-3 = Strong applicant. Outstanding scores and academic performance, clear evidence of leadership in research and/or community service, and clinical excellence.
>> Examples of a 1: Honors in all rotations; outstanding step scores; letters indicating the best student faculty have seen in their past 25 years of teaching; applicant led a new educational, research, or service program with profound impact; top-tier first-author paper or numerous high-impact publications/research program.

> 5 = Average applicant. Published a few papers, good scores and solid grades that are mostly honors. This could be someone who is very strong in 1 domain but has weaknesses in other domains, or someone who has moderate strength in all domains.

> 7-9 = Below-average applicant. This could be someone who has moderate weaknesses in all domains, or lacks strength in all domains.

Medical School Rankings:

Medical school ranking was determined via four alternative approaches (eTable 3): top 10 school per US News and World report, top 20 school per US News and World report, school affiliated with a top 10

ophthalmology hospital per US News and World report, and school affiliated with a top 10 ophthalmology program per Ophthalmology Times (best overall, best clinical, best research, and/or best residency).[1-3]

Reviewer Faculty Lines and Scope:

Faculty rank was classified as none/clinical instructor, assistant, associate, or full professor. Faculty line was classified as clinician educator, medical center line, or affiliated/adjunct. At Stanford, clinician educator faculty are focused on clinical work and education with some academic scholarship (frequently publish but are not expected to do so); medical center line faculty perform a hybrid of clinical work and research/scholarship with variable percent effort devoted to each, and are expected to publish as well as to teach; and affiliated/adjunct faculty are predominantly clinical, either part-time or with the majority of their effort at one of the affiliated teaching sites (Veterans Affairs or county hospital). No faculty reviewers of residency applications were in the research-focused university tenure line.

Statistical Analysis and Regression Modeling:

We only used one variable for medical school ranking due to collinearity. Top 20 medical school (US News and World Report) was selected because it had the strongest association with scores (t-test statistic). AΩA status was simplified to a binary variable (not elected versus all others). MD versus DO medical degree was omitted from the final model due to small sample size (9 applicants after missing data exclusions), unlikely to have high influence on model fits based on residuals with our best-fit model (eFigure 3). We similarly did not include reviewer URiM status due to insufficient sample size (only 2 URiM reviewers). Since faculty type variables demonstrated collinearity, only faculty rank was used (stronger association with scores than faculty line or location). Faculty rank was evaluated as a binary variable (full professor versus other, since assistant and associate professor had approximately the same effect size). Only a subset of applicants had USMLE Step 2 scores provided at application submission; these were evaluated in a separate model.

eReferences.

1. Best medical schools: research. US News and World Report. Accessed August 16, 2020. 2 https://www.usnews.com/best-graduate-schools/top-medical-schools/research-rankings

2. Best hospitals for ophthalmology. US News and World Report. Accessed August 16, 2020. https://health.usnews.com/best-hospitals/rankings/ophthalmology

3. Charters L. Bascom Palmer tops in annual program survey. Ophthalmology Times. Accessed August 16, 2020. https://www.ophthalmologytimes.com/view/bascom-palmer-tops-annual-program-survey

eTable 1. Parsimonious Best-Fit Multivariable Linear Regression Model of Applicant and Reviewer Characteristics on Preliminary Application Scores, using 2019-2020 Residency Application Data[1,2,3]

	Beta coefficient	95% Confidence Interval	p-value[4]
Applicant Characteristics			
Top 20 medical school (USNews and World Report)[5]	-1.15	-1.44 to -0.86	<0.0001
2nd Degree (e.g., PhD, MS, MPH)	-0.71	-1.07 to -0.36	<0.0001
Alpha Omega Alpha[6]			
Not elected (vs all others)	0.57	0.26 to 0.89	0.0004
Step 1 score (10-point increments)	-0.39	-0.48 to -0.30	<0.0001
Reviewer Characteristics			
Faculty Rank as Full Professor	0.60	0.12 to 1.08	0.02

URiM = Underrepresented in Medicine

[1] Model performed with crossed random effects for applicants and reviewers

[2] Excluded three applicants with missing Step 1 scores (osteopathic students).

[3] Reviewer URiM status and allopathic versus osteopathic medical degree variable omitted from model due to small sample size

[4] p-values calculated based on χ^2-test comparison of (a) this multivariable model without any interactions vs. (b) the same model except omitting the specified characteristic.

[5] Used top 20 medical school (USNews and World Report) criterion for medical school ranking. This ranking had stronger effect size than any of the other rankings, which were consequently dropped due to collinearity.

[6] Incomplete data since not all schools have Alpha Omega Alpha and many students submitted applications before Alpha Omega Alpha elections at their medical schools

eTable 2. Parsimonious Best-Fit Multivariable Linear Regression Model of Applicant and Reviewer Characteristics on Preliminary Application Scores, using 2019-2020 Residency Application Data among Applicants with Step 2 Scores[1,2,3]

	Beta coefficient	95% Confidence Interval	p-value[4]
Applicant Characteristics			
Top 20 medical school (USNews and World Report)[5]	-0.83	-1.39 to -0.26	0.005
2nd Degree (e.g., PhD, MS, MPH)	-0.47	-1.01 to 0.08	0.10
Alpha Omega Alpha[6] (reference = not yet determined)			
Not elected (vs all others)	0.48	0.01 to 0.94	0.04
Step 1 score (10-point increments)	-0.29	-0.46 to -0.11	0.001
Step 2 score (10-point increments)	-0.30	-0.50 to -0.09	0.005
Reviewer Characteristics			
Faculty Rank as Full Professor	0.93	0.42 to 1.45	0.0006

URiM = Underrepresented in Medicine

[1] Model performed with crossed random effects for applicants and reviewers

[2] Excluded three applicants with missing Step 1 scores (osteopathic students).

[3] Reviewer URiM status and allopathic versus osteopathic medical degree variable omitted from model due to small sample size

[4] p-values calculated based on χ^2-test comparison of (a) this multivariable model without any interactions vs. (b) the same model except omitting the specified characteristic.

[5] Used top 20 medical school (USNews and World Report) criterion for medical school ranking. This ranking had stronger effect size than any of the other rankings, which were consequently dropped due to collinearity.

[6] Incomplete data since not all schools have Alpha Omega Alpha and many students submitted applications before Alpha Omega Alpha elections at their medical schools

eTable 3. Medical School Rankings

|	
Top 10 medical school (USNews and World Report	Harvard University
	Johns Hopkins University
	University of Pennsylvania
	New York University
	Stanford University
	Columbia University
	Mayo Clinic
	University of California Los Angeles
	University of California San Francisco
	Washington University St Louis
Top 20 medical school (USNews and World Report)	Harvard University
	Johns Hopkins University
	University of Pennsylvania
	New York University
	Stanford University
	Columbia University
	Mayo Clinic
	University of California Los Angeles
	University of California San Francisco
	Washington University St Louis
	Cornell University
	Duke University
	University of Washington
	University of Pittsburgh
	University of Michigan
	Yale University
	University of Chicago
	Northwestern University
	Vanderbilt University
	Icahn School of Medicine at Mt. Sinai
Top 10 ophthalmology hospital (USNews and World Report)	University of Miami Bascom Palmer Eye Institute
	Thomas Jefferson University Wills Eye Hospital
	Johns Hopkins University Wilmer Eye Institute
	Harvard University Massachusetts Eye and Ear Infirmary
	University of California Los Angeles Jules Stein Eye Institute
	Iowa University
	Duke University
	University of Michigan Kellogg Eye Institute
	University of California San Francisco
	Cleveland Clinic Cole Eye Institute
	University of Miami Bascom Palmer Eye Institute
	Thomas Jefferson University Wills Eye Hospital

Top 10 ophthalmology program (Ophthalmology Times – best overall, best research, best residency, and/or best clinical)	Johns Hopkins University Wilmer Eye Institute
	Harvard University Massachusetts Eye and Ear Infirmary
	University of Michigan Kellogg Eye Institute
	Iowa University
	Duke University
	University of Utah Moran Eye Center
	University of California Los Angeles Jules Stein Eye Institute
	Oregon Health Sciences University Casey Eye Institute
	University of California San Francisco
	Medical University of South Carolina Storm Eye Institute

eFigure 1. 2D Histogram of Redacted and Unredacted Preliminary Residency Application Rankings, 2019-2020
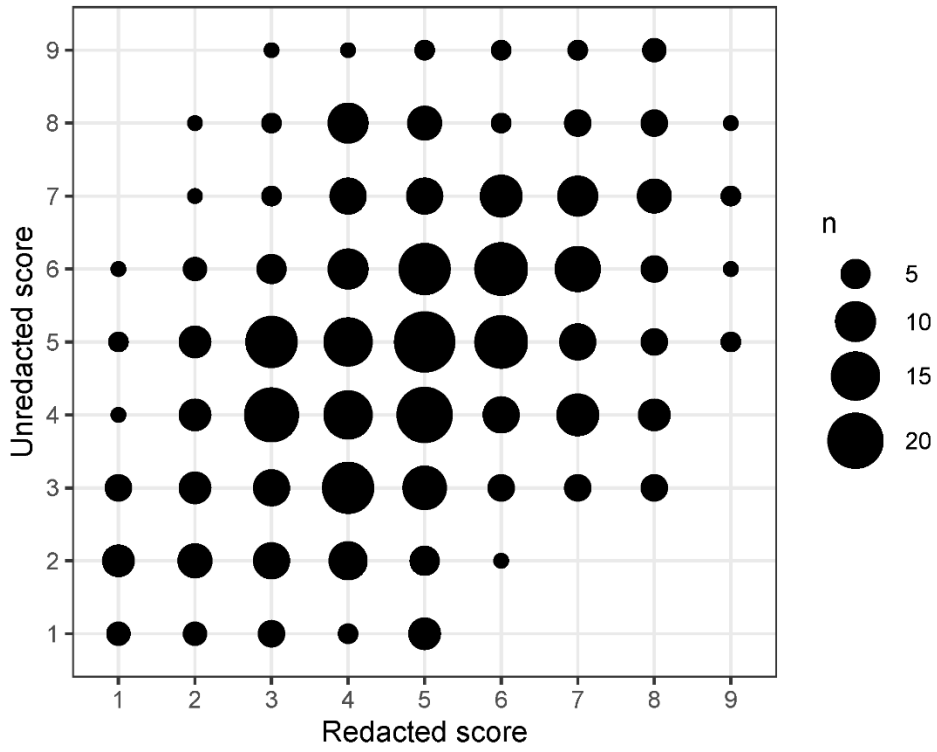
eFigure 2. Distribution of Preliminary Residency Application Rankings based on International Medical Graduate Status, 2019-2020
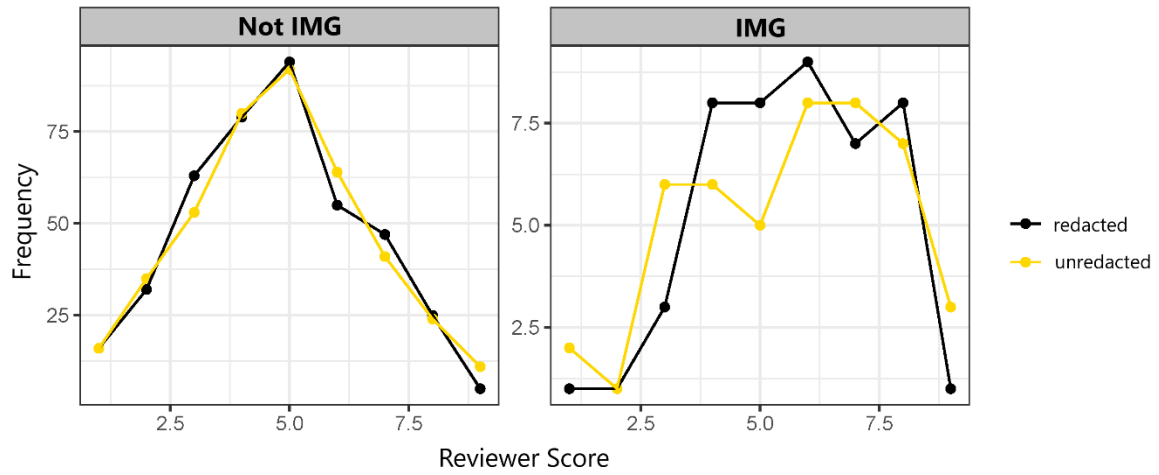
Distribution of redacted and unredacted reviewer scores among International Medical Graduate (IMG) and non-IMG applicants. p = 0.89 (IMG) and 0.51 (non-IMG) for a statistically-significant difference between redacted and unredacted scores.

eFigure 3. Residency Application Ranking Best-Fit Model Residuals for Osteopathic Applicants