

Bias in Radiology Resident Selection: Do We Discriminate Against the Obese and Unattractive?

Charles M. Maxfield, MD, Matthew P. Thorpe, MD, PhD, Terry S. Desser, MD, Darel E. Heitkamp, MD, Nathan C. Hull, MD, Karen S. Johnson, MD, Nicholas A. Koontz, MD, Gary W. Mlady, MD, Timothy J. Welch, MD, and Lars J. Grimm, MD

Abstract

Purpose

To evaluate for appearance-based discrimination in the selection of radiology residents.

Method

A deception study simulating the resident selection process examined the impact of attractiveness and obesity on resident selection. Seventy-four core faculty from 5 academic radiology departments reviewed mock residency applications in September and October 2017. Each application included demographic information and a photograph, representing a prespecified distribution of facial attractiveness and obesity, combined with randomized academic and supporting variables. Reviewers

independently scored applications for interview desirability. Reviewer scores and application variables were compared using linear mixed fixed- and random-effects models.

Results

Reviewers evaluated 5,447 applications (mean: 74 applications per reviewer). United States Medical Licensing Examination Step 1 scores were the strongest predictor of reviewer rating ($B = 0.35$ [standard error (SE) = 0.029]). Applicant facial attractiveness strongly predicted rating (attractive vs unattractive, $B = 0.30$ [SE = 0.056]; neutral vs unattractive, $B = 0.13$ [SE = 0.028]). Less influential but still significant predictors included race/

ethnicity ($B = 0.25$ [SE = 0.059]), preclinical class rank ($B = 0.25$ [SE = 0.040]), clinical clerkship grades ($B = 0.23$ [SE = 0.034]), Alpha Omega Alpha membership ($B = 0.21$ [SE = 0.032]), and obesity (vs not obese) ($B = -0.14$ [SE = 0.024]).

Conclusions

Findings provide preliminary evidence of discrimination against facially unattractive and obese applicants in radiology resident selection. Obesity and attractiveness were as influential in applicant selection for interview as traditional medical school performance metrics. Selection committees should invoke strategies to detect and manage appearance-based bias.

The tendency to attribute traits and abilities to individuals based on their physical appearance is well documented.^{1,2} It starts early in life^{3,4} and is seen across cultures.⁵ Unattractive people are perceived to be less intelligent, less socially skilled, and less successful than more attractive individuals.^{1,6} Obese individuals are vulnerable to similar stereotypes, including perceptions that they are unmotivated, undisciplined, and unintelligent.⁷ Such stereotypes can lead to prejudice and discrimination, as has been demonstrated in social,⁸ medical,⁹ legal,¹⁰ political,^{11,12} and occupational^{13,14} contexts.

The effects of appearance-based bias have been extensively studied in the occupational domain, where it has been shown that unattractive individuals, compared with more attractive individuals, are less likely to be hired and are offered lower starting salaries, among several negative job-related outcomes.¹⁴ Empirical evidence of weight-based discrimination in the workplace is even more robust: The obese encounter disadvantages in hiring decisions,¹³ compensation,¹⁵ and promotion.¹⁶

It might be expected that a similar bias would affect the admissions/selection process in higher education. While obesity (but not facial unattractiveness) has been shown to be associated with lower educational attainment,¹⁷ and obese high school students are less likely to attend college,¹⁸ there is little empiric evidence to attribute this to weight-based discrimination in the admissions process. Nor have any studies, to our knowledge, found evidence of discrimination in higher education admissions based on facial attractiveness.

The requirement for a photograph in graduate medical education (GME)

applications introduces the potential for bias based on physical attributes¹⁹ and presents an opportunity to study the impact of applicants' physical appearance on the selection process. In this study, we carried out a simulated resident selection process in which core faculty at 5 academic radiology departments reviewed and scored fictitious residency applications, believing they were evaluating actual applicants as part of their department's resident selection process. Our goal was to evaluate for appearance-based discrimination in the selection of GME residents to radiology residency.

Method

Subjects

Volunteers were solicited from the core faculty of 5 geographically diverse academic radiology departments (Duke University, University of Indiana, Mayo Clinic, University of New Mexico, and Stanford University) to review applications under the guise of resident application screening. This is the same pool of faculty from which our programs draw each year when help is needed in applicant screening and

Please see the end of this article for information about the authors.

Correspondence should be addressed to Charles M. Maxfield, Department of Radiology, 1905 Children's Health Center, Box 3808, Duke University Medical Center, Durham, NC 27710; telephone: (919) 684-7585; email: charles.maxfield@duke.edu.

Acad Med. 2019;94:1774–1780.

First published online May 28, 2019

doi: 10.1097/ACM.0000000000002813

Copyright © 2019 by the Association of American Medical Colleges

An AM Round blog post on this report is available at academicmedicineblog.org.

interviews, but faculty involved in actual resident application screening during the concurrent application cycle were excluded from the study. Demographic summary data aside from gender were not collected for participating faculty, to ensure anonymity given the sensitive nature of the research question.

Ethics statement

The institutional review boards (IRBs) at the 5 participating institutions exempted or approved this study, which used deception of subjects. The 3 institutions that approved the study granted waiver of consent.

Applications

We created mock applications to model the Electronic Residency Application Service (ERAS).²⁰ Each application consisted of 1 of 76 fixed baseline identities, anchored by the application photograph, combined with randomized academic and supporting variables (Table 1). The 76 baseline identities were fixed with a distribution of gender, race/ethnicity, facial attractiveness, and obesity variables, which was chosen to reflect the distribution of actual radiology applications, but with overrepresentation of certain groups to maximize statistical power. We randomized academic variables important in the selection of radiology applicants for interview (preclinical class rank, clinical clerkship grades, Alpha Omega Alpha [AOA] Honor Society membership, and quantity of research publications)^{21,22} for each application and reviewer, such that each reviewer saw a different combination of academic variables associated with any given photograph. Additional supporting variables (common volunteer activities and characteristic premedical accomplishments), deemed noninfluential in the selection process (by 3 experienced residency program directors), were randomized to each application to increase realism. Personal statements, Medical School Performance Evaluations (MSPE, Dean's Letter), letters of recommendation, specific publication citations, United States Medical Licensing Exam (USMLE) Step 2 scores, and additional advanced degrees were not included in these abbreviated applications.

Photographs

We standardized 170 open-access/stock color photographs from the Internet into the format typical for

Table 1

Definitions of Study Variables Used in Simulated Resident Selection at 5 Academic Radiology Departments, From a Study of Potential Bias Against Obese and Unattractive Applicants, 2017

Variable	Description
Baseline identities	Fixed variables assigned to each of the 76 applications provided to all reviewers
Photograph	Open-access/stock images from the Internet chosen to present prespecified distribution of gender (53% male, 47% female), race/ethnicity (35% white, 32% Asian, 29% black, 4% Hispanic), facial attractiveness (22% more attractive, 43% neutral, 34% less attractive), and obesity (52% not obese, 48% obese)
Name	Ethnically appropriate name randomly generated by behindthename.com ²³
Date of birth	Randomly generated between January 1, 1987, and December 31, 1991
Place of birth	Randomly generated U.S. city
Undergraduate school	Randomly generated university based in the United States
College major	Randomly generated science-based major of study
Academic variables	Variables deemed important selection criteria that were randomly generated uniquely for each reviewer
Preclinical quartile ranking	First, second, or third; approximately 33% chance of assignment to each category
Clinical clerkship grades	Pass, high pass, or honors in internal medicine, surgery, obstetrics–gynecology, pediatrics, family medicine, and psychiatry; weighted to give approximately 50% honors, 40% high pass, and 10% pass
Grade in radiology elective	Pass, honors, or pass/fail; weighted to give approximately 40% honors, 20% pass, and 40% “pass (pass/fail)”
AOA membership	Yes or no: AOA status was correlated to preclinical quartile ranking such that first-quartile applicants were 30% likely to be AOA, second-quartile applicants were 15% likely to be AOA, and third-quartile applicants were not eligible to be AOA
USMLE Step 1 score	Randomly generated score along a wide bell curve distribution from 215 to 270
Research publications	Total number published (mean, 1.6; range, 0–4) and submitted (1.0; 0–2)
Abstract/poster presentations	Total number at local (mean, 1.1; range, 0–3) or national (0.5; 0–1) meetings
Medical school	A medical school randomly generated from among the top 80 of the <i>U.S. News and World Report Best Medical Schools</i> for 2018, ⁴¹ excluding schools affiliated with participating sites
Supporting variables	Variables deemed not important in the selection process but included to increase realism of the application
Volunteer activities	One to three items chosen from a list of noninfluencing volunteer activities (e.g., secretary, radiology interest group; coordinator, university health clinic; mentor, West End Senior Center)
Premedical awards	One to three items chosen from a list of common and noninfluencing awards (e.g., Dean's Scholar, Order of the Key Society, Magna Cum Laude Merit Award)

Abbreviations: AOA indicates Alpha Omega Alpha Honor Society; USMLE, United States Medical Licensing Examination.

residency applications, featuring full front view of the head and shoulders of a professionally dressed individual. Photographs featuring a range of attractiveness and obesity, varying by gender and race/ethnicity, were sought. No photographs contained any identifying information, and no photographs were familiar to any of the reviewers. A panel of 8 radiologists, chosen to reflect the demographic distribution of the core radiology faculty

of the 5 departments (4 male, 4 female; 5 white, 2 Asian, 1 African American; age range, 26–56), subjectively rated the obesity and facial attractiveness of each photograph. Obesity was rated from 0 to 2 (0 = not obese; 1 = mildly obese; 2 = very obese). Facial attractiveness was rated from 1 to 5 (1 = extremely unattractive; 2 = unattractive; 3 = neutral; 4 = attractive; 5 = extremely attractive). We selected photographs based on mean ratings and narrow

interrater variability. Intraclass correlation coefficient for attractiveness was 0.73 (95% confidence interval, 0.68, 0.78); for obesity, it was 0.87 (0.84, 0.90). The final 76 photographs were selected to acquire the desired diversity of gender (53% male, 47% female), race/ethnicity (35% white, 32% Asian, 29% black, 4% Hispanic), facial attractiveness (22% more attractive, 43% neutral, 34% less attractive), and obesity (52% not obese, 48% obese). For analysis, applicants were binned into 3 attractiveness groups: “less attractive,” “neutral,” and “more attractive,” and into obese and nonobese groups using natural breaks in the data (i.e., nadirs of the histogram between peaks in trinomial distributions of attractiveness and obesity), with further binning of attractive and obese groups together. Preliminary analysis validated that effect estimates and correlation coefficients were similar using these bins compared with continuous mean attractiveness and obesity. Accordingly, we present data using bins for simplicity.

Platform

The 76 baseline identities that comprised the application pool were presented to reviewers through a “.org” website built using PHP version 5.6.30 (Zend Technologies, Cupertino, California) and MySQL version 5.1.73 (Oracle Corporation, Redwood Shores, California) deployed through a commercial shared server on hostgator.com. We designed the website to convincingly model the ERAS website. A unique, randomly generated 6-character identifier, encoded only for site and reviewer gender, was embedded in a URL provided to each reviewer by a site-specific investigator (T.S.D., D.E.H., C.M.M., G.W.M., T.J.W.). On following the link, this identifier was encrypted via a one-way salted hashing algorithm through the PHP CRYPT_SHA512 function. Only the resulting one-way hash was stored in the database, such that ratings from reviewers following their original, unique link could be obtained over multiple sessions; however, reviewer identities could not be decrypted by the website developer/statistical analyst (M.P.T.) to protect reviewers’ anonymity. The website randomly generated academic variables unique for each reviewer. These algorithm-generated variables were stored in the database, along with the reviewer’s ratings of each applicant.

Procedure

We carried out the experiment in September and October 2017, consistent with the pretense that reviewers were contributing to their department’s actual concurrent resident application screening process. The 5 site-specific investigators (1 current program director, 1 immediately past program director, 2 vice chairs of education, and 1 departmental chairman) sent emails to each member of their core faculty seeking volunteers to screen residency applications. Those faculty who responded were sent a link to the reviewer website, which presented the 76 applications in random order. Reviewers were told that applications were abbreviated for efficient review, but contained information sufficient for screening. Each was asked to review at least 50 applications and to score each application from 1 (“least desirable for interview”) to 5 (“most desirable for interview”). We told reviewers that their score would be 1 of 2 or 3 generated for each applicant, with the cumulative score used to determine interview decisions. To encourage a holistic approach, reviewers were given minimal instructions and provided only with a benchmark range of USMLE Step 1 scores (230–260) described as “typical for our program.” We asked reviewers not to discuss individual applicants until the process was complete and assured them their scores would be confidential. After completion of the study, all participants were debriefed in accordance with requirements of the IRBs of the 5 participating sites.

Statistical analysis

Statistical analysis was performed using the R statistical programming language version 3.4.3, including lme4 package version 1.1.14. (R Foundation for Statistical Computing, Vienna, Austria). We modeled applicant ratings using linear mixed effects, with random intercepts to mitigate ceiling and floor effects resulting from an individual reviewer tending to rate applicants closer to 5 or closer to 1, and random slopes to account for heterogeneous individual reviewer value placed on various parameters. The initial model included 5-way interaction terms for reviewer gender, applicant gender, applicant race/ethnicity, obesity, and attractiveness. Only the obesity–attractiveness interaction was significant; other interaction terms

were removed from the final analysis one at a time in backwards fashion based on the Bayesian information criterion. On preliminary analysis, scores for black and Hispanic applicants trended together, so we combined these as “underrepresented minorities,” given the proportionately smaller sample of Hispanic base identities. For the purposes of analysis, published and submitted peer-reviewed manuscripts had similar influence and were combined; poster presentations were not influential, and we omitted these from the final statistical model. Cumulative performance on core clinical clerkships was quantified as the sum of 0 for each pass, 1 for each high pass, and 2 for each honors grade.

Results

Of the 90 core faculty at the 5 institutions who responded to a request for volunteers, 74 followed the provided link and reviewed the mock applications (range of 7–30 reviewers at each institution). Reviewers (37 female, 37 male) evaluated an average of 74 applications (range, 23–76); 88% ($n = 65$) completed all 76 applications and 97% ($n = 72$) completed at least 75% of applications. On a scale from 1 to 5, reviewers gave applicants a mean score of 3.5 (standard deviation [SD] = 1.0), with mild leftward skewness of -0.4 . The average rating given by each reviewer ranged from 2.2 to 4.4, and the random intercept term modeling reviewer-specific ratings explained 17% of the total variability in ratings. Mixed-model adjusted values were normally distributed with mean 3.0 (SD = 0.4), without significant residual error across reviewers. No institutional differences in ratings were demonstrated.

Table 2 reports demographic distribution of the 5,447 randomly generated residency applications that were reviewed. Table 3 demonstrates the relative influence of academic and nonacademic variables on reviewer ratings. USMLE Step 1 score was the strongest predictor of ratings, with a 10-point change in Step 1 score predicting a change of 0.35 in the reviewer-adjusted rating. Expressed as a standardized regression coefficient, a 1-SD increase in Step 1 score predicted a reviewer-adjusted rating 1.2 SDs higher. The applicant’s facial attractiveness strongly predicted ratings for attractive versus unattractive

Table 2

Demographic Distribution of 5,447 Randomly Generated Residency Applications Reviewed in Simulated Radiology Resident Selection Experiment at 5 Academic Radiology Departments, From a Study of Potential Bias Against Obese and Unattractive Applicants, 2017

Characteristic	White		Asian		Black ^a		Hispanic ^a		Total
	F	M	F	M	F	M	F	M	
Facial attractiveness									
Less attractive									
Not obese	143	215	142	142	143	146	0	0	931
Obese	146	215	142	145	144	141	0	0	933
Neutral									
Not obese	142	140	143	142	142	72	0	71	852
Obese	144	287	284	290	142	289	70	0	1,506
More attractive									
Not obese	142	216	144	144	74	215	73	0	1,008
Obese	143	0	0	0	74	0	0	0	217
Total	860	1,073	855	863	719	863	143	71	5,447

Abbreviations: F indicates female; M, male.

^aBlack and Hispanic applicants were combined for analysis.

($B = 0.30$ [standard error (SE) = 0.056]) and neutral versus unattractive ($B = 0.13$ [SE = 0.028]). Applicant race was strongly associated with ratings. Reviewers at participating institutions preferred black and Hispanic applicants relative to white ($B = 0.25$ [SE = 0.059]) or Asian ($B = 0.28$ [SE = 0.023]) applicants. There were no significant interactions of race with gender, obesity, or attractiveness. Traditional medical school performance metrics were predictive of reviewer scores, including preclinical class rank ($B = 0.25$ [SE = 0.040] for first vs third quartile), clinical clerkship grades ($B = 0.23$ [SE = 0.034] for top vs lowest tertile), and AOA membership ($B = 0.21$ [SE = 0.033]). Obese applicants received lower scores compared with otherwise equivalent nonobese applicants ($B = -0.14$ [SE = 0.024]).

Collectively, the randomized academic variables (USMLE Step 1 score, clinical clerkship grades, preclinical class rank, AOA membership, and total number of research publications) accounted for 34% of the variability in ratings (Table 3). Race/ethnicity explained 6% of the variability in ratings, and the physical appearance of the applicant (facial attractiveness and obesity) explained 5%.

Figure 1 depicts reviewer-adjusted ratings for each combination of weight and attractiveness. The interaction term for obesity * attractiveness was -0.03 ($P <$

$.001$), indicating that obese applicants derived less benefit from being facially attractive than did nonobese applicants. Figure 1 also depicts the proportion of applicants in each category of obesity and attractiveness falling into the bottom 85% of applicants, an empirical threshold based on the participating residency programs interviewing approximately 15% of applicants. According to this analysis, an attractive, nonobese applicant is 14% more likely to be invited for an interview than is an equivalent unattractive, obese applicant. Figure 1 also demonstrates how obesity neutralizes much of the benefit of facial attractiveness: Obese applicants who were facially attractive were only slightly more likely to clear the 85th percentile threshold for interview than were obese applicants who were less attractive. The benefit of facial attractiveness was significantly greater for nonobese applicants.

Discussion

Our findings demonstrate significant relationships between the physical appearance of applicants and the decision to grant interviews in the setting of GME. Across the spectrum of race, gender, and academic achievement, there was a clear pattern of discrimination against facially unattractive and obese applicants.

In our simulated residency selection process involving core faculty from

5 different radiology programs, the facial attractiveness of an applicant, as presented by the application photograph, was more influential in selection for interview than were well-established medical school performance metrics such as preclinical class rank, clinical clerkship grades, AOA membership, and quantity of research publications, all of which have been shown to be among the most important factors in the selection of radiology applicants for interview.²³ While not as influential as facial attractiveness, obesity was on par with most academic metrics. Furthermore, our findings demonstrate a statistically significant interaction between facial attractiveness and obesity, such that, assuming a 15% threshold for recommendation for interview (the average of the 5 participating residency programs), an applicant who is obese and facially unattractive is 14% less likely to receive an interview than is an applicant who is nonobese and facially attractive, according to our model.

Our study builds on similar work in the business literature demonstrating improved hiring rates for attractive individuals, but few studies have explored this phenomenon in the higher education admissions process. No published studies, to our knowledge, have demonstrated discrimination in admissions based on the facial attractiveness of applicants, but 2 studies found evidence for weight-based discrimination. In 1966, Canning and Mayer showed that obese students were less likely to be accepted into elite colleges,²⁴ and more recently, Burmeister and colleagues demonstrated that applicants with a higher body mass index were less likely to be offered a position in a graduate psychology program.²⁵

The design of our study, which simulated the actual resident selection process through deception of the application reviewers, allowed us to control for several confounding variables present in prior studies. By using a single photograph as a surrogate for attractiveness, we were able to isolate the physical features of the applicant and eliminate confounders that might otherwise manifest in the admissions process. In Burmeister and colleagues' study, personal interviews were used to evaluate the influence of applicant obesity on admissions decisions, but this methodology failed to control for

Table 3

Relative Influence of Academic and Nonacademic Variables on Faculty Reviewer Ratings of Mock Residency Applicants at 5 Academic Radiology Departments, From a Study of Potential Bias Against Obese and Unattractive Applicants, 2017

Predictor	B, regression coefficient for predicted change in rating of 1-5 (standard error)	ΔR^2m , change in marginal R^2 (% change) from full model ^{a,c}	ΔR^2c , change in conditional R^2 (% change) from full model ^{b,c}
Reviewer-specific baseline (random intercept)	—	—	0.17
USMLE Step 1 score, each 10 points	0.35 (0.029)	0.26	0.28
Facial attractiveness		0.01	0.04
Attractive vs unattractive	0.30 (0.056)		
Neutral vs unattractive	0.13 (0.028)		
Race/ethnicity		0.02	0.06
URM vs Asian	0.28 (0.023)		
URM vs white	0.25 (0.059)		
White vs Asian	0.03 (0.021)		
Preclinical class rank		0.01	0.02
1st vs 3rd quartile	0.25 (0.040)		
2nd vs 3rd quartile	0.12 (0.020)		
Clinical clerkship grades		0.01	0.02
Top vs lowest tertile	0.23 (0.034)		
Middle vs lowest tertile	0.11 (0.023)		
AOA membership	0.21 (0.033)	0.01	0.01
Obesity	-0.14 (0.024)	0.01	0.01
Research publications (each manuscript published or in review)	0.06 (0.012)	0.01	0.01
Gender	0.01 (0.022)	< 0.01	0.01

Abbreviations: USMLE indicates United States Medical Licensing Examination; URM, underrepresented minority; AOA, Alpha Omega Alpha Honor Society.

^aMarginal R^2 represents explained variance by fixed effects only.

^bConditional R^2 represents explained variance by fixed and random effects together.

^c R^2 changes are derived as difference when entering each term last in a model containing all other listed variables.⁴²

indirect factors known to be correlates of obesity, such as self-confidence and interpersonal skills.²⁶ Canning’s methodology, a retrospective review of admission data, is similarly plagued by confounders that correlate with obesity and facial unattractiveness, such as letters of recommendation,²⁷ extracurricular activities,²⁶ and grades.²⁸ This problem was avoided in our study by use of fictitious applicants, the randomization of all academic variables, the exclusion of letters of recommendation, and the inclusion of only noninfluencing extracurricular activities and premedical accomplishments.

Although it was not the primary focus of the study, we also evaluated the influence of race and gender. Applicant

race was strongly influential on reviewer rating, as reviewers favored black and Hispanic applicants over white and Asian applicants. We do not consider our results inconsistent with recent studies that show implicit antiblack bias in doctors²⁹ and in a medical school admission committee,³⁰ but rather our results illustrate that implicit bias may not be a reliable predictor of behavior and should not be presumed to be a surrogate for discrimination. We suspect that our reviewers were prioritizing applicants they believed best met institutional goals and values. Regarding applicant gender, published studies in the psychology and business literature suggest that the influence of physical appearance may be stronger for female applicants than for male applicants,³¹ but the results are

inconsistent.³² We found no significant influence of applicant gender, reviewer gender, or their interaction.

We find no reason to believe that our findings are limited to radiology resident selection. Implicit antiobesity attitudes are widely held,³³ and those held by health professionals³⁴ may manifest behaviorally in clinical decision making.³⁵ Unlike in business, where physical attractiveness has been shown to correlate with success,³⁶ there is no justification in medicine for bias based on physical appearance. Resident selection committees should invoke strategies to detect and manage appearance-based bias. Existing diversity training programs should consider including, or emphasizing, education to counter appearance-based bias in their curricula. ERAS should reconsider the role of photographs in the application process.

This study has several limitations. We used deception to simulate the resident selection process, but there were differences in our applications, and in our application process, which might have resulted in our subjects behaving differently from actual reviewers of real applications. Our mock applications omitted important application factors (personal statements, letters of recommendation) for greater efficiency in review and to eliminate confounders; psychology research suggests that with less information, reviewers are more likely to rely on nonacademic factors.³⁷ Our reviewers might have weighed certain academic or nonacademic factors differently from those primarily responsible for application screening during the concurrent application cycle, or from those reviewers with more experience in the task. We attempted to compensate for this by including only core faculty with experience in evaluating residency or fellow applicants. Only volunteer activities and premedical accomplishments felt to be noninfluential by 3 experienced program directors were included in the applications, but it is possible that individual supporting variables might have influenced certain reviewers. Limited demographic data were collected for subjects, to ensure anonymity given the sensitive nature of the research question, and to maintain compliance with our waiver of consent. The attractiveness and obesity of

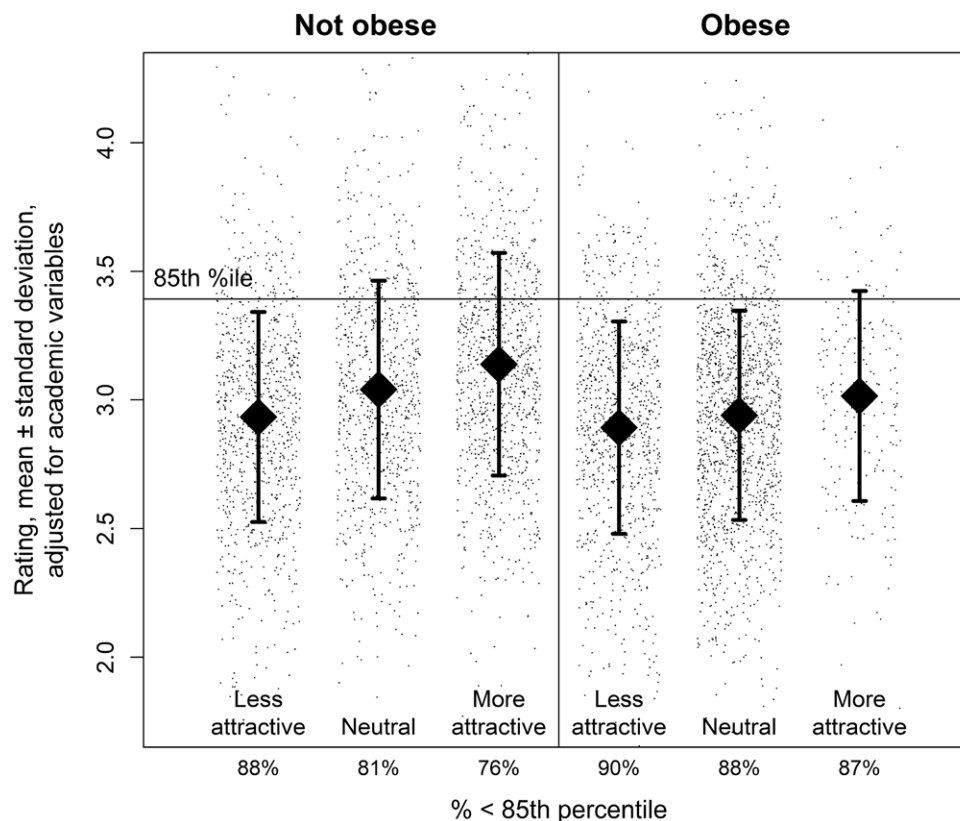


Figure 1 Distribution of application ratings from simulated residency selection process at 5 academic radiology departments, from a study of potential bias against obese and unattractive applicants, 2017. Ratings are on a scale of 1–5, by obesity and attractiveness, adjusted for individual reviewer average ratings using a mixed random-intercepts model. The horizontal line reflects the 85th percentile for ratings across all groups; values below this threshold would not be expected to receive an invitation for interview, based on our participating institutions' empirical invitation rates. Percentages along the bottom of the plot represent the group-specific rejection rates. Mean ratings and rejection rates demonstrate a significant linear benefit of attractiveness among nonobese applicants ($P < .01$ with Holm adjustment for multiple comparisons) but an attenuated, nonsignificant benefit of attractiveness among obese applicants.

applicants were determined by a panel of radiologists. A larger panel, or one with a different composition of members, might have rated photographs differently. We attempted to compensate for this by choosing applicant photographs with narrow interrater variability. Assessment of physical attractiveness is subjective and may be subject to cultural and social forces.^{38,39} Photographs were used as a static cue for obesity. This simplifies and may underestimate the influence of obesity. To maximize statistical power, certain demographic groups were over- or underrepresented in our study. Females accounted for only 29% of applications to radiology programs in 2017⁴⁰ but comprised 47% of applicants in our experiment. Blacks accounted for just 6% of applicants to radiology programs in 2017 but in our experiment accounted for 29%. Obese applicants were overrepresented in our simulated application pool, in the subjective opinion of our experienced program directors.

In conclusion, our study provides preliminary evidence for discrimination against facially unattractive and obese applicants in admissions to GME radiology residency programs. We hope these findings raise the awareness of admissions decision makers as to the potential influence of appearance-based bias. We recommend that ERAS reconsider the role of photographs in the GME application process.

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: Granted by Indiana University Institutional Review Board (IRB no. 1708886894) on September 11, 2017; Mayo Clinic IRB (no. 17-007970) on October 2, 2017; and Stanford University IRB (ref no. 42833) on September 29, 2017. Waived by Duke University (IRB Pro00086317) on August 9, 2017, and University of New Mexico Human Research Review Committee (submission ID: 17-338) on October 10, 2017.

C.M. Maxfield is vice chair of education, Department of Radiology, Duke University Medical Center, Durham, North Carolina.

M.P. Thorpe is a radiology resident, Department of Radiology, Duke University Medical Center, Durham, North Carolina.

T.S. Desser is professor, Department of Radiology, Stanford University Medical Center, Stanford, California.

D.E. Heitkamp is staff radiologist and associate residency program director, Florida Hospital, Orlando, Florida.

N.C. Hull is assistant professor, Department of Radiology, Mayo Clinic, Rochester, Minnesota.

K.S. Johnson is residency program director, Department of Radiology, Duke University Medical Center, Durham, North Carolina.

N.A. Koontz is director of fellowship programs, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, Indiana.

G.W. Mlady is chair, Department of Radiology, University of New Mexico, Albuquerque, New Mexico.

T.J. Welch is associate chair of education, Department of Radiology, Mayo Clinic, Rochester, Minnesota.

L.J. Grimm is assistant professor, Department of Radiology, Duke University Medical Center, Durham, North Carolina.

References

- 1 Thorndike EL. A constant error in psychological ratings. *J Appl Psychol.* 1920;4:25–29.
- 2 Dion K, Berscheid E, Walster E. What is beautiful is good. *J Pers Soc Psychol.* 1972;24:285–290.
- 3 Cramer P, Steinwert T. Thin is good, fat is bad: How early does it begin? *J Appl Dev Psychol.* 1998;19:429–451.
- 4 Dion KK. Young children's stereotyping of facial attractiveness. *Dev Psychol.* 1973;9:183–188.
- 5 Dion KK, Pak AW, Dion KL. Stereotyping physical attractiveness: A sociocultural perspective. *J Cross Cult Psychol.* 1990;21:158–179.
- 6 Langlois JH, Kalakanis L, Rubenstein AJ, Larson A, Hallam M, Smoot M. Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol Bull.* 2000;126:390–423.
- 7 Puhl RM, Heuer CA. The stigma of obesity: A review and update. *Obesity (Silver Spring).* 2009;17:941–964.
- 8 Gortmaker SL, Must A, Perrin JM, Sobol AM, Dietz WH. Social and economic consequences of overweight in adolescence and young adulthood. *N Engl J Med.* 1993;329:1008–1012.
- 9 Phelan SM, Burgess DJ, Yeazel MW, Hellerstedt WL, Griffin JM, van Ryn M. Impact of weight bias and stigma on quality of care and outcomes for patients with obesity. *Obes Rev.* 2015;16:319–326.
- 10 Sigall H, Ostrove N. Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. *J Pers Soc Psychol.* 1975;31:410–414.
- 11 Berggren N, Jordahl H, Poutvaara P. The looks of a winner: Beauty and electoral success. *J Public Econ.* 2010;94:8–15.
- 12 Roehling PV, Roehling M, Brennan AR, et al. Weight bias in U.S. candidate selection and election. *Equal Divers Inclus.* 2014;33:334–346.
- 13 Roehling MV. Weight-based discrimination in employment: Psychological and legal aspects. *Pers Psychol.* 1999;52:969–1016.
- 14 Hosoda M, Stone-Romero EF, Coats G. The effects of physical attractiveness on job-related outcomes: A meta-analysis of experimental studies. *Pers Psychol.* 2003;56:431–462.
- 15 Register CA, Williams DR. Wage effects of obesity among young workers. *Soc Sci Q.* 1990;71:130–141.
- 16 Bordieri JE, Drehmer DE, Taylor DW. Work life for employees with disabilities: Recommendations for promotion. *Rehabil Couns Bull.* 1997;40:181–191.
- 17 Fowler-Brown AG, Ngo LH, Phillips RS, Wee CC. Adolescent obesity and future college degree attainment. *Obesity.* 2010;18:1235–1241.
- 18 Crosnoe R. Gender, obesity, and education. *Sociol Educ.* 2007;80:241–260.
- 19 Kogan M, Frank RM. A picture is worth a thousand words: Unconscious bias in the residency application process? *Am J Orthop (Belle Mead NJ).* 2015;44:E358–E359.
- 20 Association of American Medical Colleges. Electronic Residency Application Service. <https://www.aamc.org/services/eras>. Accessed May 3, 2019.
- 21 National Resident Matching Program, Data Release and Research Committee. Results of the 2016 NRMP Program Director Survey. Washington, DC: National Resident Matching Program; 2016.
- 22 Grimm LJ, Shapiro LM, Singhapricha T, Mazurowski MA, Desser TS, Maxfield CM. Predictors of an academic career on radiology residency applications. *Acad Radiol.* 2014;21:685–690.
- 23 Behind the Name. Random name generator. <https://www.behindthename.com/random>. Accessed May 3, 2019.
- 24 Canning H, Mayer J. Obesity—Its possible effect on college acceptance. *N Engl J Med.* 1966;275:1172–1174.
- 25 Burmeister JM, Kiefner AE, Carels RA, Musher-Eizenman DR. Weight bias in graduate school admissions. *Obesity (Silver Spring).* 2013;21:918–920.
- 26 Mobius MM, Rosenblat TS. Why beauty matters. *Am Econ Rev.* 2006;96:222–235.
- 27 Nicklin JM, Roch SG. Biases influencing recommendation letter contents: Physical attractiveness and gender. *J Appl Soc Psychol.* 2008;38:3053–3074.
- 28 French MT, Robins PK, Homer JF, Tapsell LM. Effects of physical attractiveness, personality, and grooming on academic performance in high school. *Labour Econ.* 2009;16:373–382.
- 29 Sabin J, Nosek BA, Greenwald A, Rivara FP. Physicians' implicit and explicit attitudes about race by MD race, ethnicity, and gender. *J Health Care Poor Underserved.* 2009;20:896–913.
- 30 Capers Q 4th, Clinchot D, McDougle L, Greenwald AG. Implicit racial bias in medical school admissions. *Acad Med.* 2017;92:365–369.
- 31 Heilman ME, Saruwatari LR. When beauty is beastly: The effects of appearance and sex on evaluations of job applicants for managerial and nonmanagerial jobs. *Organ Behav Hum Perfor.* 1979;23:360–372.
- 32 Eagly AH, Ashmore RD, Makhijani MG, Longo LC. What is beautiful is good, but...: A meta-analytic review of research on the physical attractiveness stereotype. *Psychol Bull.* 1991;110:109.
- 33 Greenwald AG, Poehlman TA, Uhlmann EL, Banaji MR. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *J Pers Soc Psychol.* 2009;97:17–41.
- 34 Sabin JA, Marini M, Nosek BA. Implicit and explicit anti-fat bias among a large sample of medical doctors by BMI, race/ethnicity and gender. *PLoS One.* 2012;7:e48448.
- 35 Smedley BD, Stith AY, Nelson AR, eds. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington, DC: National Academies Press; 2003.
- 36 McElroy JC, DeCarlo TE. Physical attractiveness on cognitive evaluations of saleswomen's performance. *J Mark Theory Pract.* 1999;7:84–100.
- 37 Locksley A, Hepburn C, Ortiz V. Social stereotypes and judgments of individuals: An instance of the base-rate fallacy. *J Exp Soc Psychol.* 1982;18:23–42.
- 38 Thornhill R, Gangestad SW. Facial attractiveness. *Trends Cogn Sci.* 1999;3:452–460.
- 39 Cunningham MR, Roberts AR, Barbee AP, Druen PB, Wu CH. "Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *J Pers Soc Psychol.* 1995;68:261.
- 40 Association of American Medical Colleges. Statistics: Cross specialty applicant data. <https://www.aamc.org/services/eras/stats/359278/stats.html>. Accessed May 3, 2019.

References cited only in the tables

- 41 U.S. News and World Report. 2019 best medical schools. <https://www.usnews.com/best-graduate-schools/top-medical-schools?int=a4d609>. Accessed May 15, 2019.
- 42 Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol Evol.* 2013;4:133–142.