

# ACADEMIC MEDICINE

Journal of the Association of American Medical Colleges

Uncomposed, edited manuscript published online ahead of print.

This published ahead-of-print manuscript is not the final version of this article, but it may be cited and shared publicly.

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

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

**DOI:** 10.1097/ACM.0000000000002813

ACCEPTED

**Academic Medicine**

**DOI: 10.1097/ACM.0000000000002813**

**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

**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 a 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.

Correspondence should be addressed to Charles 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.

*Funding/Support:* None reported.

*Other disclosures:* None reported.

*Ethical approval:* Granted by Indiana University Institutional Review Board (IRB# 1708886894) on September 11, 2017; Mayo Clinic IRB (#17-007970) on October 2, 2017; and Stanford University IRB (Ref #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.

## **Abstract**

### **Purpose**

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

### **Method**

A deception study simulating the resident selection process to examine the impact of attractiveness and obesity on resident selection. Seventy-four core faculty from five academic radiology departments reviewed mock residency applications in September and October 2017. Applications included demographic information and 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 versus unattractive,  $B = 0.30$  [SE = 0.056]; neutral versus 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 (versus 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.

ACCEPTED

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

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

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,<sup>17</sup> and obese high school students are less likely to attend college,<sup>18</sup> 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 attributes<sup>19</sup> and presents an opportunity to study the impact of the applicant's physical appearance on the selection process. In this study, we carried out a simulated resident selection process in which core faculty at five 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 five 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 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, in order to ensure anonymity given the sensitive nature of the research question.

### **Ethics statement**

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

### **Applications**

We created mock applications to model the Electronic Residency Application Service (ERAS).<sup>20</sup> Each application consisted of one 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)<sup>21,22</sup> 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 non-influential in the selection process (by three 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 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 eight radiologists, chosen to reflect the demographic distribution of the core radiology faculty of the five departments (four male, four female; five white, two Asian, one African American; age range 26–56), subjectively rated the obesity and facial attractiveness of each photograph. Obesity was rated from 0 – 2 (0 = not obese; 1 = mildly obese; 2 = very obese). Facial attractiveness was

rated from 1–5 (1 = extremely unattractive; 2 = unattractive; 3 = neutral; 4 = attractive; 5 = extremely attractive). We selected photographs based on mean ratings and narrow inter-rater variability. Intraclass correlation coefficient for attractiveness was 0.73 (95% confidence interval [CI] 0.68, 0.78); for obesity, 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 non-obese 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 to 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, CA) and MySQL version 5.1.73 (Oracle Corporation, Redwood Shores, CA) 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 five site-specific investigators (one current program director, one immediately past program director, two vice chairs of education, and one 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 one of two or three 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 IRB of the five 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” (URM), 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 five 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 applications); 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 modelling 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 .35 in the reviewer adjusted rating.

Expressed as a standardized regression coefficient, a 1 standard deviation increase in Step 1 score predicted a reviewer adjusted rating 1.2 standard deviations higher. The applicant's facial attractiveness strongly predicted ratings for attractive versus unattractive (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 1st versus 3rd quartile), clinical clerkship grades (B = 0.23 [SE = 0.034] for top versus lowest tertile), and AOA membership (B = 0.21 [SE = 0.033]). Obese applicants received lower scores compared to otherwise equivalent non-obese 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 obese applicants derived less benefit from being facially attractive than did non-obese 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, non-obese 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 non-obese applicants.

## **Discussion**

Our findings demonstrate significant relationships between the physical appearance of applicants and the decision to grant interviews in the setting of graduate medical education. 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 five 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.<sup>23</sup> 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 five 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 non-obese and facially attractive, according to our model.

Our study builds upon 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 two 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,<sup>24</sup> 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.<sup>25</sup>

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 indirect factors known to be correlates of obesity, such as self-confidence and interpersonal skills.<sup>26</sup> 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,<sup>27</sup> extracurricular activities,<sup>26</sup> and grades.<sup>28</sup> 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 non-influencing 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 anti-black bias in doctors<sup>29</sup> and in a medical school admission committee,<sup>30</sup> but rather 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 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,<sup>31</sup> but the results are inconsistent.<sup>32</sup> We found no significant influence of applicant gender, reviewer gender, or their interaction.

We find no reason to believe our findings are limited to radiology resident selection. Implicit anti-obesity attitudes are widely held,<sup>33</sup> and those held by health professionals<sup>34</sup> may manifest behaviorally in clinical decision-making.<sup>35</sup> Unlike in business, where physical attractiveness has been shown to correlate with success,<sup>36</sup> 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. The National Resident Matching Program (NRMP) 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 than 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 non-academic factors.<sup>37</sup> Our reviewers might have weighed certain academic or nonacademic factors differently than those primarily responsible for application screening during the concurrent application cycle, or than 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 three 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, in order 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 applicants was 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 inter-rater variability. Assessment of physical attractiveness is subjective and may be subject to cultural and social forces.<sup>38,39</sup>

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 2017<sup>40</sup> 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 the NRMP reconsider the role of photographs in the GME application process.

## References

1. Thorndike EL. A constant error in psychological ratings. *J Appl Psychol.* 1920;4:25–29.
2. Dion KK, 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.* 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, Ryn VM. 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. *Equality, Diversity and Inclusion: An International Journal*. 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*. 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*. 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 Economics*. 2009;16:373–382.
29. Sabin, DJ, Nosek DB, Greenwald DA, Rivara DF. Physicians’ implicit and explicit attitudes about race by MD race, ethnicity, and gender. *J Health Care Poor Underserved*. 2009;20:896.
30. Capers Q, Clinchot D, McDougale 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. 2012;7:48448.
35. Smedley BD, Stith AY, Nelson AR, eds. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care.* Washington, DC: The National Academies Press; 2003.
36. McElroy JC, DeCarlo TE. Physical attractiveness on cognitive evaluations of saleswomen's performance. *Journal of Marketing Theory and Practice.* 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. ERAS 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. 2019 Best Medical Schools. U.S. News and World report. <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 R<sup>2</sup> from generalized linear mixed-effects models. *Methods Ecol Evol.* 2013;4:133–142.

## Figure Legend

### Figure 1

Distribution of application ratings from simulated residency selection process at five 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 non-obese applicants ( $P < .01$  with Holm adjustment for multiple comparisons) but an attenuated, nonsignificant benefit of attractiveness among obese applicants.

**Table 1****Definitions of Study Variables Used in Simulated Resident Selection at Five Academic Radiology Departments, From a Study of Potential Bias Against Obese and Unattractive Applicants, 2017**

Variable	Description
<b>Baseline identities</b>	Fixed variables assigned to each of the 76 applications provided to all reviewers
Photograph	Open access/stock images from 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 <sup>23</sup>
Date of birth	Randomly generated between January 1, 1987, and December 31, 1991
Place of birth	Randomly generated United States city
Undergraduate school	Randomly generated university based in the United States
College major	Randomly generated science-based major of study
<b>Academic variables</b>	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 and gynecology, pediatrics, family medicine and psychiatry; weighted to give approximately 50% honors, 40% high pass, 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 1st quartile applicants were 30% likely to be AOA, 2nd quartile applicants were 15% likely to be AOA, and 3rd 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 U.S. News and World Report Best Medical Schools for 2018, <sup>41</sup> excluding schools affiliated with participating sites

<b>Supporting variables</b>	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 non-influencing volunteer activities (e.g., secretary, radiology interest group; coordinator, university health clinic; mentor, West End Senior Center)
Pre-medical awards	One to three items chosen from a list of common and non-influencing 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.

**Table 2**

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

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

Abbreviations: F indicates female; M, male.

<sup>a</sup>Black and Hispanic applicants were combined for analysis.

**Table 3**

**Relative Influence of Academic and Nonacademic Variables on Faculty Reviewer Ratings of Mock Residency Applicants at Five 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)	$\Delta R^2_m$ , change in marginal $R^2$ (% change) from full model <sup>a,c</sup>	$\Delta R^2_c$ , change in conditional $R^2$ (% change) from full model <sup>b,c</sup>
<b>Reviewer specific baseline (random intercept)</b>	-	--	0.17
<b>USMLE Step 1 score, each 10 points</b>	0.35 (0.029)	0.26	0.28
<b>Facial attractiveness</b>		0.01	0.04
Attractive vs. unattractive	0.30 (0.056)		
Neutral vs. unattractive	0.13 (0.028)		
<b>Race/ethnicity</b>		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)		
<b>Preclinical class rank</b>		0.01	0.02
1st vs. 3rd quartile	0.25 (0.040)		
2nd vs. 3rd quartile	0.12 (0.020)		
<b>Clinical clerkship grades</b>		0.01	0.02
Top vs. lowest tertile	0.23 (0.034)		
Middle vs. lowest tertile	0.11 (0.023)		
<b>AOA membership</b>	0.21 (0.033)	0.01	0.01
<b>Obesity</b>	-0.14 (0.024)	0.01	0.01
<b>Research publications (each manuscript published or in review)</b>	0.06 (0.012)	0.01	0.01
<b>Gender</b>	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.

<sup>a</sup>Marginal  $R^2$  represents explained variance by fixed effects only.

<sup>b</sup>Conditional  $R^2$  represents explained variance by fixed and random effects together.

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

Figure 1

