

Patient Representation in Medical Literature: Are We Appropriately Depicting Diversity?

Jonathan P. Massie, MD*
 Daniel Y. Cho, MD, PhD†
 Cameron J. Kneib, BS†
 Jacob R. Burns, MD, MBA‡
 Christopher S. Crowe, MD†
 Megan Lane, MD§
 Afaaf Shakir, MD¶
 Danielle L. Sobol, MD†
 Janice Sabin, PhD, MSW||
 Janelle D. Sousa, MD†
 Eduardo D. Rodriguez, MD,
 DDS**
 Thomas Satterwhite, MD††
 Shane D. Morrison, MD, MS†

Background: Racial disparities exist in the accessibility, delivery, and quality of healthcare. Clinical images are central to plastic surgery, but choice of images in the literature is susceptible to implicit biases. The objective of this study was to determine if published images reflect the racial demographic of patients.

Methods: A search for color photographs and rendered graphics depicting human skin was completed in 6 plastic surgery journals and the *New England Journal of Medicine Images in Clinical Medicine* for each decade between 1992 and 2017. For each article, images were categorized as white or nonwhite based on Fitzpatrick Scale (1–3 versus 4–6). Additionally, the authors' geographic region was documented. Proportional data and average number of nonwhite images per article were compared. Regression analyses were performed to assess the correlation of time and geographic region on nonwhite images.

Results: In total, 24,209 color photographs and 1,671 color graphics were analyzed. In plastic surgery journals, 22% of photographs were nonwhite and the average number of photographs per article with white skin was 5.4 compared with 1.6 with nonwhite skin ($P < 0.0001$). There was a significant increase in nonwhite photographs over time ($r = 0.086$, $P < 0.001$) and association of nonwhite photographs with international authors ($r = 0.12$, $P < 0.001$).

Conclusions: Roughly 60%–70% of the world population and 30% of US cosmetic patients are nonwhite. Images in plastic surgery literature reflect neither racial demographics by global region nor the patient population seeking surgery. To advance equitable care, images should better represent the racial composition of the populations served. (*Plast Reconstr Surg Glob Open* 2019;7:e2563; doi: 10.1097/GOX.0000000000002563; Published online 26 December 2019.)

INTRODUCTION

Medical journals are a primary source of continuing

From the *Division of Plastic Surgery, Department of Surgery, Feinberg School of Medicine at Northwestern University, Chicago, Illinois; †Division of Plastic Surgery, Department of Surgery, University of Washington School of Medicine, Seattle, Washington; ‡Department of Surgery, University of California Davis Medical Center, Sacramento, California; §Section of Plastic Surgery, Department of Surgery, University of Michigan School of Medicine, Ann Arbor, Michigan; ¶Division of Plastic Surgery, Department of Surgery, University of Chicago Medicine, Chicago, Illinois; ||Department of Biomedical Informatics and Medical Education, Center for Health Equity, Diversity, and Inclusion, University of Washington, Seattle, Washington; **Hansjörg Wyss Department of Plastic Surgery, New York University School of Medicine, New York, New York; and ††Align Surgical Associates, San Francisco, California.

Received for publication August 30, 2019; accepted October 16, 2019.

J.P.M., D.Y.C., and C.J.K. contributed equally to this work.

Copyright © 2019 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of The American Society of Plastic Surgeons. This is an open-access article distributed under the terms of the [Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 \(CCBY-NC-ND\)](#), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

DOI: 10.1097/GOX.0000000000002563

medical education. Previous work has shown that racial and ethnic bias is ubiquitous in medical education in textbooks, Microsoft PowerPoint (Microsoft Corp., Redmond, WA, USA) slides, and simulated clinical scenarios.^{1–4} Furthermore, non-conscious racial and ethnic biases exist among physicians and others with doctoral degrees and may contribute to racial and ethnic healthcare disparities.⁵ These biases impact patient provider communication, treatment decisions, and specific areas of patient care such as pain management.^{6–8}

In many specialties, there is a reliance on photographs and images to depict diagnoses, injury patterns, and interventions. In particular, plastic surgery relies heavily on the depiction of human skin for demonstration of the clinical problems treated. These photographs and images are often in color and patient skin tone is discernable. Both authors and journals drive the decision of which images are included in medical literature, and it is possible that unconscious bias affects these decisions. Previous research has shown that implicit bias can percolate into educational material, which may then be absorbed downstream by learners.^{9,10} Racial and ethnic biases are an example of implicit biases that may be unconsciously integrated into curricula. Presumably, such bias may impact the images

Disclosure: The authors have no financial interest to declare in relation to the content of this article.

chosen for publication resulting in disparate representation of race and ethnicity within the medical literature.

Medical images also likely impact the learner at both the conscious and subconscious level. In one study, exposure to an unknown face for one-tenth of a second was enough time to make judgment.¹¹ Medical images are the cornerstone of teaching in much of plastic surgery and have a high impact on the learner. Racially biased educational materials may limit physicians' ability to identify and treat disease in patients of color. When using imagery as a teaching tool, there is an obligation to accurately reflect the racial composition of the patient population to avoid unintentionally promoting implicit racism.

The aim of this study was to determine if published medical images reflect the racial demographic of patients seeking plastic surgery. In addition, a secondary aim was to establish if published images in plastic surgery literature accurately reflected racial demographics of the general population, both within the United States and globally. The *New England Journal of Medicine* (Images in Clinical Medicine) was searched as a non-plastic surgery comparison to better generalize any findings to other medical specialties.

METHODS

Journal Selection

Six leading journals in the field of plastic surgery were selected to evaluate images: *Annals of Plastic Surgery* (APS), *Aesthetic Surgery Journal* (ASJ), *Journal of Craniofacial Surgery* (JCFS), *Journal of Hand Surgery* (JHS), *Journal of Plastic, Reconstructive, and Aesthetic Surgery* (JPRAS), and *Plastic and Reconstructive Surgery* (PRS). All articles were extracted from each journal for 2016 and the first 2 months of 2017 as this was the time when data collection started. Then, starting from 2010, all the articles were collected at 10-year intervals until the first issue the journal published figures in color was reached. For example, information from *ASJ* was collected from the years 2016 and first 2 months of 2017, 2010, 2000, and 1996, as 1996 was the first year images were consistently published in color while 2000 and 2010 were separated by a decade and 2016 along with the first 2 months of 2017 were in the year data collection started. *The New England Journal of Medicine* (NEJM) "Images in Clinical Medicine" was used as a non-plastic surgery comparison and all photographs were analyzed from every article from

1992 to April 2017. The years the articles were collected from each journal are shown in [Table 1](#).

Data Collection

Nine researchers (D.Y.C, C.J.K, J.P.M, J.R.B, C.S.C, M.L, A.S., D.L.S, and S.D.M.) collected the data by downloading all articles from the selected journals at the selected timepoints and analyzing each in its entirety for color photographs, rendered graphics, or illustrations (excluding supplementary material). Data collection criteria were established before collection and are defined below in *Categorizing Race/Ethnicity*. Ten percent of the data were rechecked and <5% were recategorized. Additional data collected from each article included the authors' country of origin. Articles were assigned to 1 of 6 geographic regions based on the United Nations classification: Africa, Americas, Asia, Europe, Oceania, or "cross regional."¹² Articles were assigned the term "cross-regional" if the authors' countries of origin spanned 2 or more geographic regions. Articles were assigned the term "international" if they originated outside of the United States.

Images

Photographs and rendered graphics depicting the human form were included in this study. Images were excluded if they did not have clear depictions of skin, including mucosal surfaces, internal anatomy, or animal studies. The context of the articles was not taken into account when categorizing the figures.

Categorizing Race/Ethnicity

Race is social construct, not a biological attribute. For the purpose of this study that focuses on medical imagery, we chose to use skin tone as a proxy for race/ethnicity. Coders categorized photographs or graphics depicting human skin into either "white" (Fitzpatrick skin ratings 1–3) or "nonwhite" (Fitzpatrick skin ratings 4–6) ([Fig. 1](#)).¹³ Although the Fitzpatrick scale is an imperfect scale for categorizing race, this has been used previously for such reasons and was a means of allowing objective categorization.¹

Interrater Reliability

Interrater reliability was determined based on a collection of 22 images chosen by an individual researcher (D.Y.C) and sent to all researchers involved in data collection for image categorization. These data were then

Table 1. White and Nonwhite Photographs and Rendered Graphics in Medical Literature

	Years Collected	Articles (n)	International Articles (n)	White Photographs (n)	Nonwhite Photographs (n)	White Graphics (n)	Nonwhite Graphics (n)
All		4,100	2,644	18,792	5,417	1,585	86
APS	2013, 2016, 2017 (January–February)	317	210	2,648	331	133	2
ASJ	1996, 2000, 2010, 2016, 2017 (January–February)	352	104	2,759	691	459	5
JCFS	2010, 2016, 2017 (January–March)	700	598	2,064	2,178	100	11
JHS	2007, 2010, 2016, 2017 (January–February)	320	143	1,349	140	98	11
JPRAS	2000, 2010, 2016, 2017 (January–February)	709	664	2,963	1,373	170	15
PRS	1998, 2000, 2010, 2016, 2017 (January–February)	870	416	5,901	461	625	42
NEJM	1992–2016, 2017 (January–April)	832	509	1,108	243	N/A	N/A

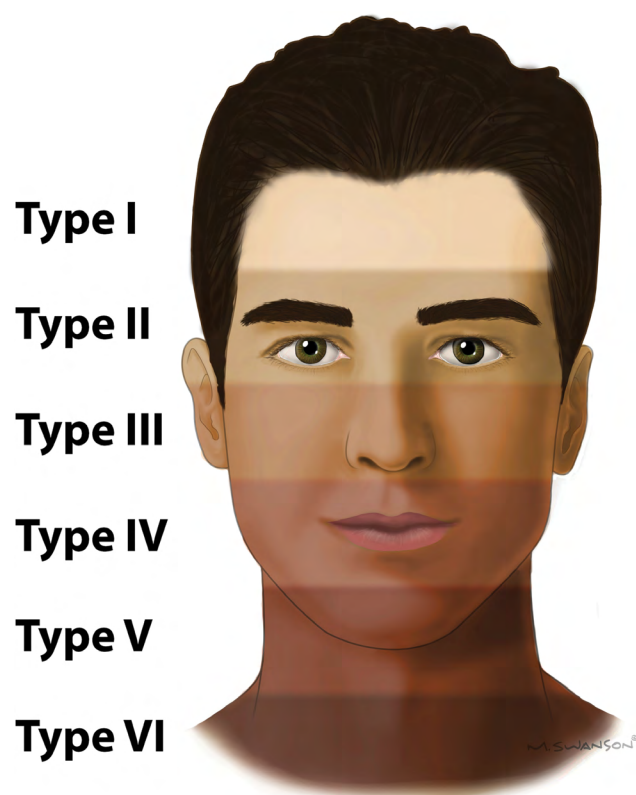


Fig. 1. Fitzpatrick Skin Scale. The Fitzpatrick skin scale numerically categorizes human skin tone into 6 categories based on response to UV light with Type I being the palest and most likely to burn with UV exposure and Type VI being deeply pigmented dark brown and least likely to burn with UV exposure.

evaluated for interrater reliability using Fleiss' kappa. Interpretation of kappa was based on Landis and Koch cutoffs for correlation reliability: 0.01–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–0.99, almost perfect agreement.

Statistical Analysis

Data were analyzed using Prism GraphPad (Version 7, La Jolla, CA, USA.) and IBM SPSS Statistical Software (Version 21, Armonk, NY, USA). The average number of white and nonwhite images per article was calculated and comparisons were made using a two-tailed unpaired Student's *t*-test. Proportional data of white versus nonwhite images were also calculated and compared per article. Univariate regression analyses comparing the average number of white and nonwhite images over time were performed. In addition, univariate regression of the proportion of nonwhite images over time was performed. Pearson's correlation coefficient (*r*) was reported for all univariate regression analyses. Finally, multivariable regression analysis was performed to control for the effect of both time and international articles on the publication of nonwhite images. Standardized coefficients (β) were reported for all multivariable regression

analyses. For all statistical analyses, significance allowed for a type I error of $\alpha = 0.05$.

RESULTS

A total of 4,100 articles were identified that had at least 1 color photograph or 1 color rendered graphic depicting human skin. Of papers analyzed, 1,456 were published from authors in the United States, while 2,644 were published from internationally based authors. A total of 24,209 individual color photographs depicting skin were analyzed, while 1,671 color rendered graphics were analyzed (Table 1). Interrater reliability for grading skin type was determined to have a Fleiss' Kappa coefficient of 0.65 ($P < 0.001$), indicating substantial agreement between raters.

Of articles reviewed, 78% of photographs (18,792 photographs) were of white skin and 22% of photographs (5,417 photographs) were of nonwhite skin (Table 1, Fig. 2A). In plastic surgery journals, the average number of photographs per article with white skin was 5.4 compared with 1.6 photographs with nonwhite skin (white versus nonwhite skin, $P < 0.0001$) (Fig. 2B). The NEJM articles in comparison had an average of 1.3 photographs per article with white skin and 0.29 photographs per article with nonwhite skin (white vs nonwhite skin, $P < 0.0001$) (Fig. 2B).

Color rendered graphics depicting human skin in plastic surgery journals were also analyzed. Ninety-five percent of graphics (1,585 graphics) were of white skin, and only 5% of graphics (86 graphics) were of nonwhite skin. The average number of white graphics per article was 0.48 compared with an average of 0.03 nonwhite graphics per article (white versus nonwhite skin, $P < 0.0001$) (Table 1, Fig. 2C). Graphics in NEJM were not analyzed for comparison as NEJM did not have any published rendered graphics.

Over the last 20 years, there was an increase in the number of nonwhite photographs represented in journals. Univariate linear regression analysis of the percentage of nonwhite color photographs showed a significant increase over time in plastic surgery journals ($r = 0.92$, $P = 0.02$), compared with a non-significant increase in the percentage of nonwhite photos in the NEJM ($r = 0.31$, $P = 0.09$) (Fig. 2D).

A large portion of nonwhite photographs in plastic surgery journals were found in papers published by international authors. US articles on average had 0.85 nonwhite photographs, whereas international articles had an average of 2.0 nonwhite photographs per article (white vs nonwhite, $P < 0.0001$) (Fig. 3A). In the NEJM, there was not a significant difference in the average number of nonwhite photographs published from US versus international authors (US average = 0.31 vs international average = 0.28, $P = 0.48$) (Fig. 3B). Regions with the overall highest percentage of nonwhite photographs in plastic surgery journals were articles from Asia (42%) and Africa (56%). Articles from the Americas (14%), Europe (9%), Oceania (11%), and cross-regional papers (14%) had a substantially lower percentage of nonwhite photographs (Fig. 3C).

When controlling for both time and international articles on multivariable regression analysis, there was a significant increase in nonwhite photographs in plastic

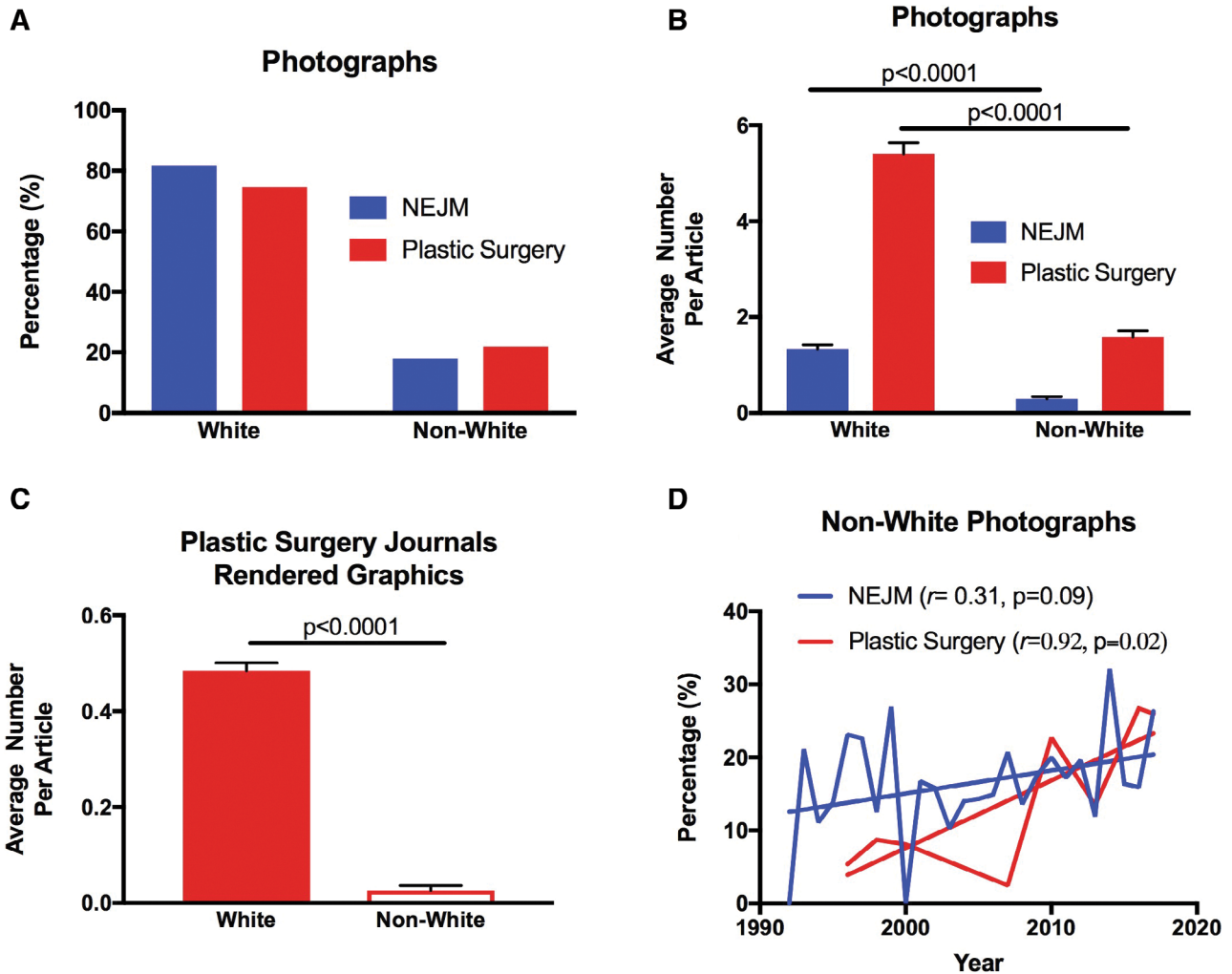


Fig. 2. A, Percentage of white and nonwhite photographs in the New England Journal of Medicine “Images in Clinical Medicine” compared with plastic surgery journals. B, Average number of white vs nonwhite photographs per article in the NEJM compared with plastic surgery journals. C, Average number of white vs nonwhite rendered graphics per article in plastic surgery journals. D, Percentage of nonwhite photographs in the NEJM and in plastic surgery journals over time.

surgery journals over time ($\beta = 0.086, P < 0.001$) and a significant association of nonwhite photographs with international articles ($\beta = 0.12, P < 0.001$) (Table 2). On the other hand, for photographs in NEJM, there was not a significant increase in nonwhite photographs over time ($\beta = 0.06, P = 0.10$) nor a significant association of nonwhite photographs with international articles ($\beta = -0.041, P = 0.26$) (Table 2).

DISCUSSION

This is the first study to examine whether racial biases exist in published surgical images in the plastic surgery literature. Plastic surgery focuses on the management and reconstruction of disease processes affecting human skin and soft tissue. Therefore, medical images reflect an important clinical decision-making tool and representation of patient outcomes. The NEJM was also searched as a comparison in an effort to generalize findings to other medical disciplines.

More than 25,000 images in the form of human photographs and rendered graphics were examined, spanning decades of medical publication. Overall, 78% of photographs depicted subjects with white skin tones (specifically Fitzpatrick skin ratings 1–3). According to estimates from the United States Census Bureau, India, Africa, and China compose about 50% of the world’s population.¹⁴ When considering other equatorial regions with similar indigenous skin tones, the world’s nonwhite population likely approaches or exceeds 60%–70%.¹⁵ Therefore, our findings suggest that the plastic surgery literature vastly overrepresents white patients at a global level. Within the United States, 23% of the population is estimated to be nonwhite.¹⁶ Our results found that from publications originating in the Americas, only 14% of photographs were of nonwhite patients. Thus, on both an international and granular scale, the plastic surgical literature does not accurately reflect the racial demographic comprising these regions.

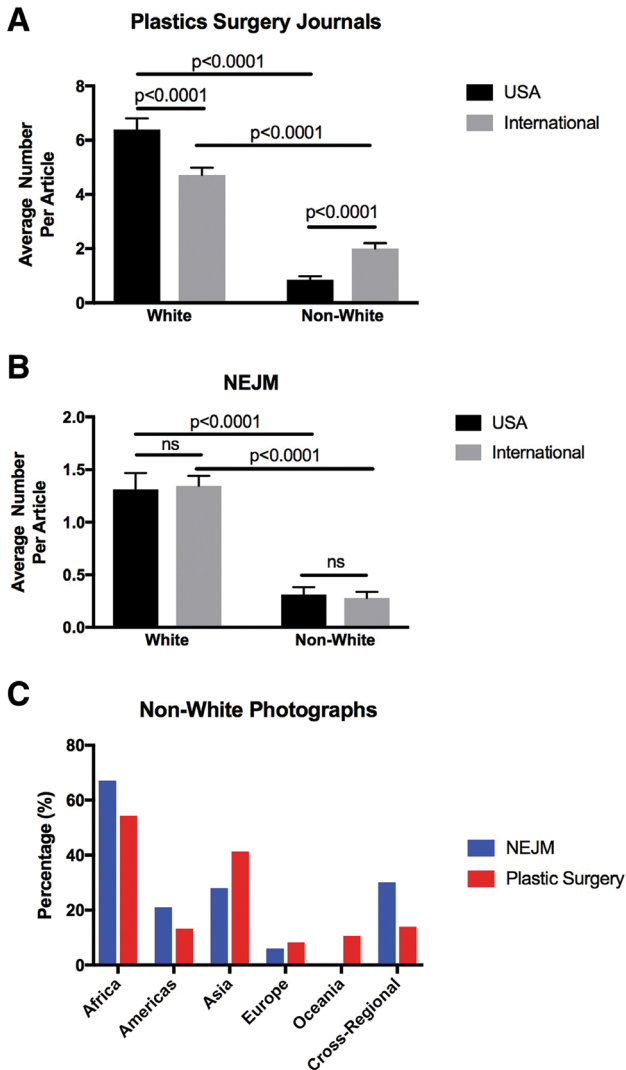


Fig. 3. A, Average number of white and nonwhite photographs per article in US-based vs International-based author groups in the combined plastic surgery journals. B, Average number of white vs nonwhite photographs per article in US-based vs international-based author groups in the NEJM. C, Percentage of nonwhite photographs in the NEJM compared with plastic surgery journals by global region of authorship.

However, the racial profile of a given world region may not match the patient demographic seeking plastic surgery. According to the American Society of Plastic Surgeons, in 2010 and 2017, 30% of cosmetic plastic surgery procedures were performed on nonwhite patients.^{17,18} In a study of trauma centers in Pennsylvania, about 15% of the overall trauma population was black.¹⁹ Therefore, when accounting for additional reconstructive procedures that occur in large volume trauma centers, the overall nonwhite patient population undergoing plastic surgery within the United States likely is in the 20%–30% range. With 86% of photographs in the plastic surgery literature originating from the Americas depicting white patients, publishing does not reflect the racial demographic of patients undergoing plastic surgery. Therefore, images in

Table 2. Association of Having Nonwhite Photographs Controlling for Year and International Articles on Multivariable Regression

	Year (β , P)	International (β , P)
All plastic surgery journals	0.086, $P < 0.001$	0.12, $P < 0.001$
APS	-0.056, $P = 0.59$	-0.125, $P = 0.72$
ASJ	0.13, $P = 0.025$	0.22, $P < 0.001$
JCFS	0.197, $P = 0.002$	1.8, $P = 0.001$
JHS	0.17, $P = 0.003$	0.066, $P = 0.23$
JPRAS	0.037, $P = 0.33$	0.007, $P = 0.85$
PRS	0.028, $P = 0.42$	-0.039, $P = 0.26$
NEJM	0.06, $P = 0.10$	-0.041, $P = 0.26$

the plastic surgery literature neither reflect racial demographics by geographic region nor the patient population seeking plastic surgery.

Trends over time do suggest that the plastic surgery literature is beginning to better reflect racial diversity. There was a statistically significant increase in the percentage of nonwhite photographs depicted in plastic surgery literature over time on both univariate and multivariate linear regression. In 2017, 25% of photographs published in plastic surgery journals were of nonwhite patients. These numbers begin to approach the 30% margin of expected nonwhite plastic surgery procedures in the United States.

In an effort to extrapolate these findings outside of the plastic surgery literature, we examined images in the NEJM. However, an inherent limitation to this study was the forum by which NEJM published clinical images in medicine. Typically, no more than 4–6 images were displayed per article, whereas it was not unusual for plastic surgery articles to portray upwards of 20 images. Therefore, comparing the average number of images per article in plastic surgery journals to NEJM is limited. Nevertheless, NEJM did display similar trends to plastic surgery journals with over 80% of all photographs depicting white skin. Unlike plastic surgery journals, there was no significant increase in nonwhite photographs over time by either univariable or multivariable linear regression. These data suggest that racial biases in medical images do not appear to be limited to the plastic surgery literature and may be pervasive throughout medical publishing.

A common theme to both plastic surgery journals and the NEJM was that studies published by authors from outside of the United States, or with cross-regional collaboration, were more likely to include nonwhite photographs. In particular, studies including authors from Africa or Asia had higher proportions of nonwhite photographs. Interestingly, even in articles originating from Africa, only 56% of photographs were of nonwhite patients.

While clinical photographs may be limited in part by the disease processes that affect certain populations, the same is not true for rendered graphics. Authors have complete control over the skin tone and implied race of rendered graphics. In this study, 95% of rendered graphics portrayed white skin. This offers an easy and particular area of future improvement for all areas of medical literature.

The work presented here has several limitations. Firstly, we made racial assumptions based on skin color

using the Fitzpatrick scale, which in effect legitimizes using an arbitrary marker to define race. Skin color is but one marker of the ethnic origin of an individual; it does not accurately reflect ancestry and can be highly variably even among siblings.²⁰ In addition, the concepts of race, ancestry, and ethnicity are convoluted since race is a social, not a scientific, construct. The idea of “race” can be traced back to a single scientist, Samuel Morton, who measured brain volumes in the 19th century.²¹ Unfortunately, Morton’s “five races” would become the basis for discrimination, bigotry, and prejudice for centuries to come. Therefore, it was not our intention to distill the patients depicted in medical images down to the color of their skin to assume race and instead was a mechanism by which to show the current biases that exist in medical publishing. Although the Fitzpatrick scale is an imperfect scale for categorizing race, this has been used previously for such reasons and was a means of allowing objective categorization.¹

The larger issue at hand then becomes how to begin to discuss and dispel racial biases that percolate from a construct that has no scientific basis. The answer, in part, is that we must first recognize and acknowledge these biases. For decades, peer-reviewed academic publications have used photographs and images that inadequately portray the diversity in demographics of patients affected by particular diseases. This is particularly striking in the lack of diversity in medical illustration. These inequities in medical reporting can have lasting downstream effects on the accessibility and provision of healthcare. However, until now, there has been no discussion of these biases within our medical peer-reviewed journals. To create change, there first needs to be mounting internal reflection to ensure we are not falling victim to our own biases. Only then can checks and balances be put into place to guarantee fair and equitable medical reporting.

In summary, we would emphasize that this study is not meant to criticize any journals or medical disciplines. Instead, we hope it serves as the crucible for a new era in medical publishing: one that places the onus on both authors and editors to reflect critically on the images that are chosen for publication. Skin tone does not define race. Ancestry does not define race. Race does not exist. But in our medical journals, we should strive for our images to reflect the global community which we serve.

Shane D. Morrison, MD, MS

Division of Plastic Surgery
University of Washington Department of Surgery
7CT73.1 Harborview Medicine Center
325 9th Avenue, Mailstop #359796
Seattle, WA 98104
E-mail: shanedm@uw.edu

ACKNOWLEDGMENT

We would like to thank Marco Swanson, MD, for his contribution to creating the medical illustration seen in Figure 1.

REFERENCES

1. Louie P, Wilkes R. Representations of race and skin tone in medical textbook imagery. *Soc Sci Med.* 2018;202:38–42.
2. Martin GC, Kirgis J, Sid E, et al. Equitable imagery in the pre-clinical medical school curriculum: findings from one medical school. *Acad Med.* 2016;91:1002–1006.
3. Tsai J, Ucik L, Baldwin N, et al. Race matters? Examining and rethinking race portrayal in preclinical medical education. *Acad Med.* 2016;91:916–920.
4. Hoffman KM, Trawalter S, Axt JR, et al. Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proc Natl Acad Sci.* 2016;113:4296–4301.
5. Sabin J, Nosek BA, Greenwald A, et al. Physicians’ implicit and explicit attitudes about race by MD race, ethnicity, and gender. *J Health Care Poor Underserved.* 2009;20:896–913.
6. Cooper LA, Roter DL, Carson KA, et al. The associations of clinicians’ implicit attitudes about race with medical visit communication and patient ratings of interpersonal care. *Am J Public Health.* 2012;102:979–987.
7. Green AR, Carney DR, Pallin DJ, et al. Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. *J Gen Intern Med.* 2007;22:1231–1238.
8. Sabin JA, Greenwald AG. The influence of implicit bias on treatment recommendations for 4 common pediatric conditions: pain, urinary tract infection, attention deficit hyperactivity disorder, and asthma. *Am J Public Health.* 2012;102:988–995.
9. Hafferty FW. Beyond curriculum reform: confronting medicine’s hidden curriculum. *Acad Med.* 1998;73:403–407.
10. Turbes S, Krebs E, Axtell S. The hidden curriculum in multicultural medical education: the role of case examples. *Acad Med.* 2002;77:209–216.
11. Willis J, Todorov A. First impressions: making up your mind after a 100-ms exposure to a face. *Psychol Sci.* 2006;17:592–598.
12. Statistics Division United Nations. *Standard country or area codes for statistical use (M49)*. New York, NY: United Nations Secretariat; 2011.
13. Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol.* 1988;124:869–871.
14. Bureau USC. U.S. and world population clock. Available at <https://www.census.gov/popclock/world>. Published 2018.
15. Jablonski NG. The evolution of human skin and skin color. *Annu Rev Anthropol.* 2004;33:585–623.
16. United States Census Bureau. U.S. Census Bureau quickFacts: United States. Available at <https://www.census.gov/quickfacts/fact/table/US/PST045217>. Published 2018.
17. American Society of Plastic Surgeons. 2010 Cosmetic Demographics Report. *Am Soc Plast Surg.* 2010.
18. American Society of Plastic Surgeons. 2017 Plastic Surgery Statistics Report. *Am Soc Plast Surg.* 2017.
19. Glance LG, Osler TM, Mukamel DB, et al. Trends in racial disparities for injured patients admitted to trauma centers. *Health Serv Res.* 2013;48:1684–1703.
20. Edmonds P. These twins will make you rethink race. *Natl Geogr Mag.* 2018.
21. Kolbert E. There’s no scientific basis for race—it’s a made-up label. *Natl Geogr Mag.* 2018.