


The Cost of Diversity: An Analysis of Representation and Cost Barriers in Stock Photo Libraries for Health Education Materials, 2021

Zachary A. Chichester, BS¹ 

Michelle A. Jewell, MS²

Catherine E. LePrevost, PhD²

Joseph G. L. Lee, PhD, MPH¹ 

Introduction. *Ineffective health communication can drive health disparities and limit the effectiveness of interventions to reduce them. Stock photo libraries are a critical tool for developers of patient education, health education, and intervention materials. It is not clear how well stock photo libraries represent communities bearing disproportionate burdens of disease. Method. We conducted a search using five popular stock image libraries (Adobe Stock Images, Canva, Getty Images, Microsoft Office Image Library, and Pixabay) in November 2021 to evaluate diversity and representation in health-related stock photos. We searched for the following five key preventive health topics: healthy eating, exercising, quitting smoking, vaccination, and pregnancy. The images (N = 495) were coded for age, gender presentation, representation of perceived minoritized racial/ethnic identity, skin color using the Massey–Martin skin color scale, markers of high socioeconomic status (SES), and access costs. Results. The representation of perceived minoritized people, darker skin color, and inclusion of markers of high SES varied greatly by the search term and library. Images predominately portrayed young adults and adults, with limited representation of other age groups. Images in libraries with any paywall were significantly more likely to depict a person of perceived minoritized racial/ethnic identity and*

depict darker skin colors, and were significantly less likely to contain markers of high SES identity than images in libraries that were free to use. Discussion. We found that it costs more to develop culturally relevant health education materials for minoritized populations and groups that do not represent high SES populations. This may hinder the development of effective communication interventions.

Keywords: stock photos; patient education; representation; material development; health promotion; equity

¹East Carolina University, Greenville, NC, USA

²North Carolina State University, Raleigh, NC, USA

Authors' Note: The authors would like to thank Abdul Zahra for assistance with establishing coding reliability. This research was supported in part by an Undergraduate Research and Creativity Award from East Carolina University. Research reported in this publication was supported by the National Library of Medicine of the National Institutes of Health under Award Number G08LM013198. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. An earlier version of this article partially fulfilled the degree requirements of the East Carolina University Honors College. Address correspondence to Zachary A. Chichester, Department of Health Education and Promotion, East Carolina University, 1000 East 5th Street Stop 529, Greenville, NC 27858, USA; e-mail: chichesterz18@students.ecu.edu.

Health Promotion Practice

Month XXXX Vol. XX, No. (X) 1–12

DOI: 10.1177/15248399221150788

Article reuse guidelines: sagepub.com/journals-permissions



© 2023 The Author(s)

Chronic disease is a leading cause of preventable illness and death, and exacerbates poor maternal health outcomes (GBD 2019 Diseases and Injuries Collaborators, 2020). Along with policies, systems, and environmental changes, effective patient education materials that promote preventive behaviors are important to mitigating morbidity and mortality. Given that one-size-fits-all interventions can exacerbate health inequities (Frohlich & Potvin, 2008), such materials can fit within efforts to ensure equity by developing segmented interventions. As Campbell and Quintilliani (2006) argued, culturally tailored¹ communications that are relevant to minoritized groups can contribute to reducing health disparities between these groups and their counterparts. Thus, it is important for practitioners to have access to resources to create culturally tailored health education materials. One critical resource for developing health promotion materials—be they websites, apps, social media content, or print material—is diverse stock photo imagery.

Examples abound of the lack of representation in health-focused visual materials. There are gaps in existing sources of imagery for health educators (Buller et al., 2015). Medical and nursing education textbooks have been shown to fail to depict diversity by race, ethnicity, and gender for both professionals and patients (Louie & Wilkes, 2018; Parker et al., 2017). Consumer facing websites for prostate cancer screening underrepresent Black and Latinx men (Loeb et al., 2022), and cancer information materials have similar deficiencies (King, 2015).

Health promotion efforts should recognize the importance of market segmentation or the creation of different messages and materials for different groups (Kreuter et al., 2014). Creating health education materials that represent the population receiving the materials likely improves the effectiveness of materials. Specifically, culturally tailored interventions have been shown to be promising (e.g., Webb Hooper et al., 2017), and tailored approaches in general may work by increasing the relevance of the message and thereby increasing engagement (Strecher et al., 2008), both of which increase the likelihood of knowledge retention or behavior change. Other researchers have shown that identifying with the people in imagery can influence behavioral intentions (Buller et al., 2018, 2019).

We investigated the diversity and representation of minoritized groups in popular stock image libraries that may be used to create health education materials. We also explored whether there are disproportionate barriers (such as paywalls and images that require a cost to use) to access images featuring representation by age,

gender presentation, perceived racial/ethnic minoritized identities, and visual markers of socioeconomic status (SES). We asked the following questions:

Research Question 1 (RQ1): How well do stock photo libraries represent the diversity of communities at risk of preventable diseases?

Research Question 2 (RQ2): Is there a correlation between the portrayal of minoritized groups and the presence of barriers (e.g., paywalls) to access the imagery?

► METHOD

To answer the research questions, we conducted a quantitative content analysis (Riffe et al., 2005) of stock photos from five stock photo libraries. We searched the five stock photo libraries—Canva, Microsoft Office Image Library, Pixabay, Adobe Stock Images, and Getty Images—in November 2021. In each stock photo library, we conducted five separate searches. We developed terms based on plain language searches, and we incorporated feedback on real-world search practices reported by two community health workers on a grant advisory panel who worked in agricultural worker outreach. Our goal in seeking this feedback was to ensure that we were approximating real-world search strategies. The following five terms relating to preventive health were used: (1) “people eating healthy,” (2) “people exercising,” (3) “people quitting smoking,” (4) “people vaccination,” and (5) “pregnant woman.” These terms were selected because of their importance to chronic disease prevention. We also included pregnancy, given the historic neglect of maternal and child health and the important role of health education in maternal health services (e.g., Women, Infants, and Children [WIC] Program nutrition education). The first 25 images displayed for each search were downloaded, given a unique identification number, and stored for coding.

We developed a codebook to translate qualitative characteristics of the imagery into quantitative data. We next briefly describe these measures; readers interested in the full codebook can find it online (supplemental file). We first coded for relevance and barriers to access. These were relevance of the image to the search topic (e.g., we excluded images of veterinary vaccination for the term “people vaccination”), presence of paywall in the library even if the individual image was not behind the paywall, and presence of a cost to access the individual image. Each of these was coded as a binary variable.

We then coded for the presence of people and their

TABLE 1
Number and Proportion of Relevant Images, November 2021, *N* = 495

	<i>Healthier diet</i>	<i>Physical activity</i>	<i>Pregnancy</i>	<i>Quitting smoking</i>	<i>Vaccination</i>
<i>Library</i>	<i>N (%)</i>				
Adobe	25 (20.7)	25 (20.7)	25 (24.5)	23 (31.9)	22 (27.8)
Canva	25 (20.7)	25 (20.7)	25 (24.5)	25 (34.7)	18 (22.8)
Getty	23 (19.0)	25 (20.7)	25 (24.5)	24 (33.3)	25 (31.6)
Microsoft	24 (19.8)	24 (19.8)	4 (3.9)	0 (0)	5 (6.3)
Pixabay	24 (19.8)	22 (18.2)	23 (22.5)	0 (0)	9 (11.4)
Total	121 (24.4)	121 (24.4)	102 (20.6)	72 (14.5)	79 (16.0)

characteristics. We coded for the presence of at least one person (cartoons and animations were not counted as people) and the presence of multiple people. We assessed characteristics based on what a reasonable person might assume to be the characteristics of the person or persons depicted. First, based on context clues and gender norms, we assessed gender presentation into the following three categories: perceived man, perceived woman, and perceived nonbinary gender. Second, we coded for perceived age, which we categorized into four groups as follows: 0—neonate, infant, and toddler; 1—school-age child, adolescent; 2—young adult, middle-aged adult; and 3—older adult. If multiple people were present in the image, the youngest and oldest individuals were coded. Age was only coded if there was a visible face in the image. Third, the coder similarly assessed whether a reasonable viewer of the image would think there was a person of a minoritized racial/ethnic identity depicted in the image. We then coded skin color, which is more easily measured than racial/ethnic identity and is itself correlated with social experiences (Klonoff & Landrine, 2000). Specifically, we used the Massey–Martin Skin Color Guide (Massey & Martin, 2003) to score and record the skin color(s) of individuals in each photo. The original scale ranges from 1 (lightest) to 10 (darkest). The scale was consolidated into ranges of 1 to 3, 4 to 6, and 7 to 10 to improve the reliability of the metric, being coded ordinally as 1, 2, and 3, respectively. If multiple people were present in an image, the lightest and darkest skin colors were recorded. Following recommendations from Hannon and DeFina (2016), we referenced the skin color chart directly while coding images. Finally, we coded for the presence of indicators of perceived high SES (e.g., quartz countertops or expensive French cookware in photos of kitchens). Any category that was not relevant to the image or that

could not be conclusively determined/was unknown for the image was coded with “99.”

To establish reliability, we used Krippendorff’s alpha (Hayes & Krippendorff, 2007). First, two coders independently coded 32 test images; reliability ranged from -0.05 to 0.93 . We discussed differences in coding and revised the coding category definitions. The two coders then coded 25 new test images and retested reliability for items with less than $\alpha = 0.75$. After the second test, Krippendorff’s alpha ranged from 0.75 to 1.0 , indicating adequate intercoder reliability.

We conducted data management in SPSS Version 27/Macintosh (IBM, Chicago, IL) and analyses in Stata IC/16.1/Macintosh (StataCorp, College Station, TX). We first calculated frequencies and percentages of responses from each coding category and stock photo site. We then implemented regression models to test the association between stock photo site and characteristics with the following three outcomes: perception as representing a minoritized racial/ethnic group, skin color, and having a marker of high SES present. We implemented the models as having a dichotomous outcome (perceived representation of a minoritized racial/ethnic group, marker of SES present) or an ordinal outcome (skin color). Because our models had some variables where combinations of characteristics had zero images, we used logistic regression models with penalized likelihood or Firth methods to reduce small-sample bias in maximum likelihood estimation (Firth, 1993) as implemented in the Stata command *firthlogit*. The East Carolina University and Medical Center institutional review board (IRB) reviewed our research protocol and determined it to be not-human-subjects research (No. 21-001983).

► RESULTS

For the study, a total of 525 stock images were

identified from the five libraries. Of these images, 495 were relevant to the topic. Table 1 shows the proportion of relevant images by topic and library. Of these relevant images, 121 (24.4%) came from the topic of healthier diet, 121 (24.4%) came from the topic of physical activity, 102 (20.6%) came from the topic of pregnancy, 72 (14.5%) came from the topic of quitting smoking, and 79 (16.0%) came from the topic of vaccination. When analyzed by library, the percentage of images containing a visible face ranged from 38.1% (Pixabay) to 63.2% (Adobe Stock Images). When analyzed by search topic, the percentage of images containing a visible face ranged from 20.0% (vaccination) to 76.6% (physical activity). In total, 49.8% of images contained a visible face.

Diversity in imagery was operationalized as gender presentation, age, perceived racial/ethnic representation, skin color, and indicators of high SES. Table 2 illustrates that, when analyzed by library and excluding photos about pregnancy, presentation of gender as a man was featured in a range of 26.3% (Getty Images) to 79.6% (Canva) of images, whereas presentation of gender as a woman was featured in a range of 46.2% (Microsoft Office Image Library) to 91.2% (Getty Images) of images. Table 3 illustrates that, when analyzed by search term, presentation of gender as a man was featured in a range of 2.1% (pregnancy) to 68.2% (vaccination) of images, whereas presentation of gender as a woman was featured in a range of 63.2% (quitting smoking) to 100% (pregnancy) of images. Overall, excluding images specific to pregnancy, 53.9% of images depicted at least one person presenting as a man and 84.6% of images depicted at least one person presenting as a woman.

Table 2 shows that, when analyzed by library, both images with one person and images with multiple people had very low representation of neonates–toddlers, children–adolescents, and older adults. A substantial majority of the photos analyzed represented young and middle-aged adults. Table 3 shows that these trends continued to be evident when photos were analyzed by search topic. In total, representation of age groups ranged from 0% (neonates–toddlers) to 87.1% (young adults–adults) of images with one person. Values for the proportion of people portrayed as the youngest person in images with multiple people ranged from 2.5% for both neonate–toddler and child–adolescent categories to 38.1% for the young adult–adult category. Values for the proportion of people portrayed as the oldest person in images with multiple people ranged from 0% (neonate–toddler, child–adolescent) to 37.5% (young adult–adult).

Variability was seen in the perceived racial/ethnic representation in stock photos when analyzed by

library in Table 2. Values ranged from 26.7% (Pixabay) to 67.7% (Getty Images) of photos featuring representation of a perceived minoritized racial/ethnic group. Variability continued to be present when analyzed by search topic in Table 3, as values ranged from 33.3% (quitting smoking) to 73.3% (vaccination) of photos featuring representation of a perceived minoritized racial/ethnic group. Overall, 51.5% of images portrayed one or more people with a perceived minoritized racial/ethnic identity.

Lighter skin colors were represented substantially more than darker skin colors in both images with one person and with multiple people. By library, Getty Images had the highest proportion of images representing darker skin colors, whereas Pixabay had the lowest proportion of images representing darker skin colors, as shown in Table 2. By search topic, vaccination had the highest proportion of images representing darker skin color, whereas quitting smoking and healthy eating featured the least representation of darker skin color, as shown in Table 3. In total, images with one person featured a person with a skin color scoring 1 to 3 on the Massey–Martin skin color scale in 77.5% of images, scoring 4 to 6 in 18.9% of images, and scoring 7 to 10 in 3.6% of images. In images with multiple people, 44.9% of images depicted the person with the lightest skin color in the image having a skin color ranging from 1 to 3, 5.1% of images depicted the person with the lightest skin color in the image having a skin color ranging from 4 to 6, and no images depicted the person with the lightest skin color in the image having a skin color ranging from 7 to 10. Continuing with images featuring multiple people, 31.3% of images depicted the person with the darkest skin color in the image having a skin color ranging from 1 to 3, 14.6% of images depicted the person with the darkest skin color in the image having a skin color ranging from 4 to 6, and only 4.1% of images depicted the person with the darkest skin color in the image having a skin color ranging from 7 to 10.

Table 2 demonstrates that, across libraries, Getty Images had the lowest proportion of images containing perceived indicators of high SES (2.5%) and Microsoft Office Image Library had the highest proportion of images containing perceived indicators of high SES (20.0%). Table 3 reveals that perceived indicators of high SES were more prevalent in the healthy eating and physical activity categories. Perceived indicators of high SES were absent in the vaccination category and minimally present in the quitting smoking and pregnancy categories. Overall, perceived indicators of high SES were present in 11.5% of images.

Table 4 presents regression models examining the association between stock photo characteristics, topics,

TABLE 2
Characteristics of Stock Photo Images by Library, November 2021, N = 495

	<i>Adobe Stock Images</i>	<i>Canva</i>	<i>Getty Images</i>	<i>Microsoft Office Image Library</i>	<i>Pixabay</i>	<i>Total</i>
<i>Characteristics coded</i>	<i>N (%)</i>					
Presence of visible face in image						
Visible face(s) present	74 (63.2)	45 (38.5)	60 (50.0)	15 (57.7)	16 (38.1)	210 (49.8)
No visible face(s)	43 (36.8)	72 (61.5)	60 (50.0)	11 (42.3)	26 (61.9)	212 (50.2)
Perceived gender presentation in image ^a						
Gender presentation as a man	45 (61.6)	39 (79.6)	15 (26.3)	8 (61.5)	5 (29.4)	112 (53.6)
Gender presentation as a woman	61 (85.9)	44 (86.3)	52 (91.2)	6 (46.2)	13 (81.3)	176 (84.6)
Gender presentation as nonbinary	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Ages represented in images with one person						
Neonate–toddler present in image	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Child, adolescent present in image	2 (5.4)	0 (0.0)	0 (0.0)	0 (0.0)	2 (16.7)	4 (3.4)
Young adult, adult present in image	32 (86.5)	15 (88.2)	37 (88.1)	7 (87.5)	10 (83.3)	101 (87.1)
Older adult present in image	3 (8.1)	2 (11.8)	5 (11.9)	1 (12.5)	0 (0.0)	11 (9.5)
Ages represented in images with multiple people						
Youngest person in image is young neonate–toddler	0 (0.0)	0 (0.0)	3 (10.0)	0 (0.0)	1 (25.0)	4 (2.5)
Oldest person in image is young neonate–toddler	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Youngest person in image is child, adolescent	1 (1.7)	0 (0.0)	1 (3.3)	2 (20.0)	0 (0.0)	4 (2.5)
Oldest person in image is child, adolescent	0 (0.0)	0 (0.0)	0 (0.0)	1 (10.0)	0 (0.0)	1 (0.6)
Youngest person in image is young adult/adult	23 (38.3)	26 (46.4)	9 (30.0)	3 (30.0)	0 (0.0)	61 (38.1)
Oldest person in image is young adult/adult	21 (35.0)	25 (44.6)	9 (30.0)	4 (40.0)	1 (25.0)	60 (37.5)
Youngest person in image is older adult	6 (10.0)	2 (3.6)	2 (6.7)	0 (0.0)	1 (25.0)	11 (6.9)
Oldest person in image is older adult	9 (15.0)	3 (5.4)	6 (20.0)	0 (0.0)	1 (25.0)	19 (11.9)
Racial/ethnic representation						
Person of minoritized racial/ethnic identity present	28 (42.4)	21 (48.8)	42 (67.7)	6 (60.0)	4 (26.7)	101 (51.5)
No person of minoritized racial/ethnic identity present	38 (57.6)	22 (51.2)	20 (32.3)	4 (40.0)	11 (73.3)	95 (48.5)
Skin color representation						
Images with one person						
Images scoring 1–3 (lightest skin color)	52 (82.5)	46 (83.6)	53 (63.9)	13 (81.3)	29 (90.6)	193 (77.5)
Images scoring 4–6	11 (17.5)	8 (14.5)	24 (28.9)	2 (12.5)	2 (6.3)	47 (18.9)
Images scoring 7–10 (darkest skin color)	0 (0.0)	1 (1.8)	6 (7.2)	1 (6.3)	1 (3.1)	9 (3.6)

(continued)

TABLE 2 (CONTINUED)

	<i>Adobe Stock Images</i>	<i>Canva</i>	<i>Getty Images</i>	<i>Microsoft Office Image Library</i>	<i>Pixabay</i>	<i>Total</i>
<i>Characteristics coded</i>	<i>N (%)</i>					
Images with multiple people						
Images scoring 1–3 (lightest skin color) for score of lightest and darkest person depicted among people present						
Lightest skin color present was 1–3	44 (44.9)	52 (49.1)	24 (36.4)	8 (50.0)	4 (50.0)	132 (44.9)
Darkest skin color present was 1–3	29 (29.6)	40 (37.8)	13 (19.7)	6 (37.5)	4 (50.0)	92 (31.3)
Images scoring 4–6 for score of lightest and darkest person depicted among people present						
Lightest skin tone present was 4–6	5 (5.1)	1 (0.9)	9 (13.6)	0 (0.0)	0 (0.0)	15 (5.1)
Darkest skin tone present was 4–6	19 (19.4)	5 (4.7)	17 (25.8)	2 (12.5)	0 (0.0)	43 (14.6)
Images scoring 7–10 (darkest skin color) for score of lightest and darkest person depicted among people present						
Lightest skin color present was 7–10	0 (0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Darkest skin color present was 7–10	1 (1.0)	8 (7.5)	3 (4.5)	0 (0.0)	0 (0.0)	12 (4.1)
Socioeconomic status						
Indicators of high socioeconomic status	19 (16.2)	14 (12.0)	3 (2.5)	5 (20.0)	7 (17.5)	48 (11.5)
No indicators of high socioeconomic status	98 (83.8)	103 (88.0)	117 (97.5)	20 (80.0)	33 (82.5)	371 (88.5)

Note. Not all coding characteristics are present for all images. Thus, numbers do not sum within each coding category to total.

^aGender presentation is calculated for all topics except pregnancy.

and libraries and the following three outcomes: perception as representing a minoritized racial/ethnic group, skin color, and having a marker of high SES present. These were predicted by the presence of faces, cost barriers to image use, and the library and topics.

Having one or more faces visible in the image was significantly associated with lower likelihood of the photo representing a perceived minoritized group (odds ratio [OR] 0.47, 95% confidence interval [CI] [0.23, 0.96]). Visible faces were not associated with skin color or having a marker of high SES shown.

Regarding cost barriers to use, having a paywall for the library was significantly associated with each outcome. Images in libraries with paywalls were significantly more likely to have representation of a perceived minoritized racial/ethnic group (OR 2.08, 95% CI [1.18, 3.66]), to have higher likelihood of a darker skin color represented (OR 1.70, 95% CI [1.09, 2.68]), and, in contrast, were less likely to have a marker of high SES included (OR 0.38, 95% CI [0.21, 0.71]).

The same pattern of results, albeit without reaching traditional thresholds of statistical significance for minoritized representation and skin color, was present based on whether the image was free to access or not.

There were also differences in outcome by library. Compared with Adobe Stock Images, Getty images were more likely to have representation of minoritized racial/ethnic groups (OR 2.80, 95% CI [1.37, 5.73]), more likely to have the presence of darker skin colors (OR 2.04, 95% CI [1.18, 3.50]), and less likely to have a marker of high SES represented (OR 0.15, 95% CI [0.05, 0.48]). Pixabay was significantly less likely than Adobe to include images with darker skin color (OR 0.25, 95% CI [0.07, 0.89]). Finally, there were differences in these three outcomes by topic. Compared with images about a healthier diet, images about vaccination had a pattern of being significantly more likely to show a perceived minoritized racial/ethnic group and darker skin color and were less likely to show a marker of high SES. Markers of high SES were

TABLE 3
Characteristics, Number, and Proportion of Images Representing the Indicated Group by Search Topic, November 2021, N = 495

	<i>Healthier diet</i>	<i>Physical activity</i>	<i>Pregnancy</i>	<i>Quitting smoking</i>	<i>Vaccination</i>	<i>Total</i>
<i>Characteristics coded</i>	<i>N (%)</i>					
Presence of visible face in image						
Visible face(s) present	50 (61.7)	82 (76.6)	50 (50.5)	15 (21.4)	13 (20.0)	210 (49.8)
No visible face(s)	31 (38.3)	25 (23.4)	49 (49.5)	55 (78.6)	52 (80.0)	212 (50.2)
Perceived gender presentation in image						
Gender presentation as a man	24 (46.2)	47 (51.1)	2 (2.1)	11 (52.4)	30 (68.2)	114 (37.6)
Gender presentation as a woman	51 (92.7)	77 (85.6)	98 (100.0)	12 (63.2)	36 (81.8)	274 (89.5)
Gender presentation as nonbinary person	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Ages represented in images with one person						
Neonate–toddler present in image	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Child, adolescent present in image	3 (12.5)	0 (0.0)	1 (2.3)	0 (0.0)	0 (0.0)	4 (3.4)
Young adult, adult present in image	21 (87.5)	28 (80.0)	43 (97.7)	9 (75.0)	0 (0.0)	101 (87.1)
Older adult present in image	0 (0.0)	7 (20.0)	0 (0.0)	3 (25.0)	1 (100)	11 (9.5)
Ages represented in images with multiple people						
Youngest person in image is young neonate–toddler	3 (5.6)	0 (0.0)	1 (16.7)	0 (0.0)	0 (0.0)	4 (2.5)
Oldest person in image is young neonate–toddler	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Youngest person in image is child, adolescent	1 (1.9)	1 (1.1)	2 (33.3)	0 (0.0)	0 (0.0)	4 (2.5)
Oldest person in image is child, adolescent	0 (0.0)	1 (1.1)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.6)
Youngest person in image is young adult/adult	18 (33.3)	37 (42.0)	0 (0.0)	1 (50.0)	5 (50.0)	61 (38.1)
Oldest person in image is young adult/adult	18 (33.3)	34 (38.6)	3 (50.0)	1 (50.0)	4 (40.0)	60 (37.5)
Youngest person in image is older adult	5 (9.3)	6 (6.8)	0 (0.0)	0 (0.0)	0 (0.0)	11 (6.9)
Oldest person in image is older adult	9 (16.7)	9 (10.2)	0 (0.0)	0 (0.0)	1 (10.0)	19 (11.9)
Racial/ethnic representation						
Person of minoritized racial/ethnic identity	22 (47.8)	33 (54.1)	19 (43.2)	5 (33.3)	22 (73.3)	101 (51.5)
No person of minoritized racial/ethnic identity	24 (52.2)	28 (45.9)	25 (56.8)	10 (66.7)	8 (26.7)	95 (48.5)
Skin color representation						
Images with one person						
Images scoring 1–3 (lightest skin color)	33 (84.6)	34 (72.3)	66 (77.6)	51 (81.0)	9 (60.0)	193 (77.5)
Images scoring 4–6	6 (15.4)	9 (19.1)	17 (20.0)	12 (19.0)	3 (20.0)	47 (18.9)
Images scoring 7–10 (darkest skin color)	0 (0.0)	4 (8.5)	2 (2.4)	0 (0.0)	3 (20.0)	9 (3.6)

(continued)

TABLE 3 (CONTINUED)

	<i>Healthier diet</i>	<i>Physical activity</i>	<i>Pregnancy</i>	<i>Quitting smoking</i>	<i>Vaccination</i>	<i>Total</i>
<i>Characteristics coded</i>	<i>N (%)</i>					
Images with multiple people						
Images scoring 1–3 (lightest skin color) for score of lightest and darkest skin color depicted among people present						
Lightest skin color present was 1–3	37 (44.0)	52 (48.1)	5 (35.7)	5 (41.7)	33 (43.4)	132 (44.9)
Darkest skin color present was 1–3	28 (33.3)	37 (34.3)	4 (28.6)	5 (41.7)	18 (23.7)	92 (31.3)
Images scoring 4–6 for score of lightest and darkest skin color depicted among people present						
Lightest skin color present was 4–6	5 (6.0)	2 (1.9)	2 (14.3)	1 (8.3)	5 (6.6)	15 (5.1)
Darkest skin color present was 4–6	7 (8.3)	13 (12.0)	3 (21.4)	1 (8.3)	19 (25.0)	43 (14.6)
Images scoring 7–10 (darkest skin color) for score of lightest and darkest skin color depicted among people present						
Lightest skin color present was 7–10	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Darkest skin color present was 7–10	7 (8.3)	4 (3.7)	0 (0.0)	0 (0.0)	1 (1.3)	12 (4.1)
Socioeconomic status						
Indicators of high socioeconomic status	15 (18.8)	26 (24.8)	6 (6.1)	1 (1.4)	0 (0.0)	48 (11.5)
No indicators of high socioeconomic status	65 (81.3)	79 (75.2)	93 (93.9)	69 (98.6)	65 (100)	371 (88.5)

also significantly less likely for pregnancy and quitting smoking compared with a healthier diet. Physical activity had a general pattern, albeit not statistically significant, of being more likely to show minoritized groups, darker skin color, and high SES than images about a healthier diet.

► DISCUSSION

Principal Findings

Our study shows that the cost to develop culturally relevant health education materials is higher for materials representing minoritized populations and populations not of high SES. Free stock images poorly represent the diversity of communities at risk of preventable diseases. Specifically, for our first research question, we found that Getty Images included more minoritized racial/ethnic groups, and that Pixabay had the least representation. Overall, photos with visible faces were less likely to represent perceived minoritized racial/ethnic groups. Getty Images also included more individuals with darker skin colors. While Pixabay had the

least representation of darker skin colors, Adobe Stock Images, Canva, and Microsoft Office Image Library also had limited representation of darker skin colors. Indeed, representation of darker skin colors was minimal overall. Getty Images also had the most representation of populations not of high SES, whereas Microsoft Office Image Library had the least representation. Perceived indicators of high SES were more common in images for healthy diet and physical activity. There was greater representation of people perceived as women than men across all topics and libraries. In addition, there was a lack of representation of children and older adults when compared with the representation of young adults/adults. For our second research question, we found that photos with cost barriers are more likely to portray minoritized groups than those without cost barriers, suggesting that there is a correlation between the portrayal of minoritized groups and the presence of barriers to access the imagery. We also found that photos with cost barriers are less likely to portray indicators of high SES than those without cost barriers.

TABLE 4
Prediction of Representation in Stock Photos in Adjusted and Unadjusted Models

<i>Predictor variable</i>	<i>Perceived minoritized racial/ethnic group represented (Y vs. N)</i>		<i>Increasing likelihood of darker skin color represented</i>		<i>Marker of high SES pictured (Y vs. N)</i>	
	<i>Unadjusted OR [95% CI]</i>	<i>Adjusted aOR [95% CI]</i>	<i>Unadjusted OR [95% CI]</i>	<i>Adjusted aOR [95% CI]</i>	<i>Unadjusted OR [95% CI]</i>	<i>Adjusted aOR [95% CI]</i>
Visible face						
No visible face	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Visible face	0.47 [0.23, 0.96]	0.57 [0.24, 1.36]	1.18 [0.76, 1.83]	1.11 [0.65, 1.91]	1.01 [0.98, 1.04]	1.04 [0.99, 1.09]
Paywall presence for at least some images in library						
No paywall	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Paywall	2.08 [1.18, 3.66]	—	1.70 [1.09, 2.68]	—	0.38 [0.21, 0.71]	—
Image is free to access						
Free to access	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Not free to access	1.58 [0.90, 2.77]	—	1.37 [0.88, 2.12]	—	0.22 [0.11, 0.47]	—
Library						
Adobe Stock Images	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Canva	1.29 [0.60, 2.77]	1.31 [0.60, 2.89]	0.75 [0.40, 1.40]	0.78 [0.41, 1.48]	0.71 [0.34, 1.47]	0.66 [0.30, 1.45]
Getty	2.80 [1.37, 5.73]	2.69 [1.30, 5.58]	2.04 [1.18, 3.50]	2.01 [1.15, 3.50]	0.15 [0.05, 0.48]	0.13 [0.04, 0.45]
Microsoft Office	1.95 [0.53, 7.13]	1.79 [0.46, 6.99]	0.73 [0.25, 2.12]	0.65 [0.21, 2.00]	1.36 [0.47, 3.91]	0.59 [0.19, 1.83]
Image Library						
Pixabay	0.53 [0.16, 1.74]	0.52 [0.15, 1.77]	0.25 [0.07, 0.89]	0.26 [0.07, 0.93]	1.13 [0.45, 2.86]	0.85 [0.31, 2.31]
Topic						
Healthier diet	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Physical activity	1.28 [0.60, 2.74]	1.34 [0.60, 2.99]	1.26 [0.65, 2.44]	1.37 [0.68, 2.76]	1.41 [0.69, 2.86]	1.42 [0.67, 3.01]
Pregnancy	0.83 [0.37, 1.89]	0.86 [0.36, 2.03]	0.88 [0.44, 1.76]	0.99 [0.48, 2.02]	0.29 [0.11, 0.77]	0.28 [0.10, 0.75]
Quitting smoking	0.57 [0.18, 1.86]	0.60 [0.18, 2.00]	0.64 [0.29, 1.41]	0.60 [0.26, 1.39]	0.09 [0.02, 0.50]	0.09 [0.02, 0.51]
Vaccination	2.88 [1.09, 7.64]	2.02 [0.68, 6.02]	2.58 [1.25, 5.29]	2.35 [1.09, 5.09]	0.03 [0.002, 0.55]	0.03 [0.002, 0.59]

Note. Number of cases in each outcome variable varies as not all images could be coded for each outcome (e.g., people presented in silhouette or images without people). Logistic regression models used penalized likelihood (i.e., Firth) methods to reduce small-sample bias in maximum likelihood estimation. Skin color is modeled in an ordinal logistic regression model. Paywall presence means the library has a paywall but not all images are necessarily behind the paywall (e.g., a library has some images behind the paywall). SES = socioeconomic status. Bold indicates significance at the $p < 0.05$ level.

Results in Context

Stock imagery is important for the development of effective health education materials. Images that do not properly represent the topic can cause patients to be misinformed, putting them at risk (Goodstein et al., 2018). Culturally relevant messaging can likely improve health outcomes (Campbell & Quintiliani, 2006) and diverse stock imagery is needed to create these materials. Previous studies have shown a lack of diversity and representation in online images depicting pregnancy, similar to the results of our study (Bogers et al., 2020). This trend has also been shown in rheumatology clinical image banks (Strait et al., 2021) and hospital social media and online platforms (Myers et al., 2019). To achieve optimal and equitable health outcomes, high-quality, highly diverse stock images need to be available to the creators of health education media.

Although racial/ethnic diversity and representation in stock imagery are important for creating culturally relevant patient education materials, materials could also take into account values, beliefs, and norms. Focusing on these “deep structures” (Resnicow et al., 1999) instead of surface-level demographics is recommended (Institute of Medicine, 2002). Such approaches can be used for personalization through “cultural tailoring” that allows interventionists to personalize based on “how individuals perceive their own culture, the extent to which they identify with it, and the specific cultural values that are important to them” (Kreuter et al., 2003, p. 137). Future research should investigate the cultural diversity and representation in stock photo libraries in combination with more tailored approaches.

It is also important to note that there are some risks to representation. For example, there is a risk that culturally specific messages around health disparities can backfire (Niederdeppe et al., 2013) by promoting fatalistic views among the intended audience and reducing risk perceptions among populations who are not represented. Researchers have documented some of these challenges in the context of messaging with cancer disparities (Landrine & Corral, 2014), HIV/STI disparities (Friedman et al., 2014), and COVID disparities (Skinner-Dorkenoo et al., 2022). Developers of materials should work to balance representation with avoiding stereotypes and perceptions that a problem is faced by just a single group.

Limitations

A limitation of this study was that, despite the establishment of intercoder reliability, perception of what qualifies as an indicator of high SES can be highly subjective.

In addition, determining whether a photo portrays a person presenting as a man or a woman, or of minoritized racial/ethnic identity, can be difficult, given that, by definition, identity comes from within. We also coded using photos from only five libraries; there are more libraries that could potentially be used by creators of health education materials. In addition, although we established intercoder reliability, this study only utilized one coder; thus it is subject to the coder's perceptions from their own identity.

Implications for Practice

Creators of health communication materials should be aware of this systemic problem and not assume that stock image libraries freely provide images that best reflect the communities they serve. At a minimum, we suggest that creators should be mindful of these library-specific limitations and allocate resources at the grant-writing stage to purchase appropriate stock photos to achieve their ultimate communication product goals. Ideally, we recommend that creators who are working with nuanced topics and specific audiences fund professional quality photo shoots to assure that appropriate, relevant, and culturally mindful images are available for their products.

Stock photo libraries should make efforts to maintain an equal pool of non-paywalled images representing minoritized and historically excluded communities. Future research in this area should investigate other groups that may feature exclusion, such as overweight/obese individuals, individuals with piercings/tattoos, and so on. Future research should also investigate the diversity and representation featured in the health communication materials themselves.

CONCLUSION

Health disparities faced by minoritized populations can be exacerbated by poor communication. Creators of health communication materials are reliant on stock photo libraries to create materials for diverse populations. The findings from this study reveal that there is substantial room for producers of free stock photos to improve their representation of minoritized populations.

ORCID iDs

Zachary A. Chichester  <https://orcid.org/0000-0002-4255-7419>
Joseph G. L. Lee  <https://orcid.org/0000-0001-9698-649X>

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available at <https://journals.sagepub.com/home/hpp>.

NOTE

1. To minimize confusion, we follow Kreuter's definition and define *tailoring* as strategies that are personalized to an individual and *targeting* as strategies that are designed to reach a population subgroup based on shared characteristics (Kreuter et al., 2003). Interventions designed for a specific cultural audience are also referred to as “culturally specific” interventions.

REFERENCES

- Bogers, L., Niederer, S., Bardelli, F., & De Gaetano, C. (2020). Confronting bias in the online representation of pregnancy. *Convergence*, 26(5–6), 1037–1059.
- Buller, M. K., Andersen, P. A., Bettinghaus, E. P., Liu, X., Slater, M. D., Henry, K., Fluharty, L., Fullmer, S., & Buller, D. B. (2018). Randomized trial evaluating targeted photographic health communication messages in three stigmatized populations: Physically-disabled, senior, and overweight/obese individuals. *Journal of Health Communication*, 23(10–11), 886–898. <https://doi.org/10.1080/10810730.2018.1536731>
- Buller, M. K., Bettinghaus, E., Buller, D. B., Liu, X., & Fluharty, L. (2015). Perceived availability of culturally and demographically diverse photographs for health education materials, Colorado, 2010. *Preventing Chronic Disease*, 12, E24. <https://doi.org/10.5888/pcd12.140450>
- Buller, M. K., Bettinghaus, E. P., Fluharty, L., Andersen, P. A., Slater, M. D., Henry, K. L., Liu, X., Fullmer, S., & Buller, D. B. (2019). Improving health communication with photographic images that increase identification in three minority populations. *Health Education Research*, 34(2), 145–158. <https://doi.org/10.1093/her/cyy054>
- Campbell, M. K., & Quintiliani, L. M. (2006). Tailored interventions in public health: Where does tailoring fit in interventions to reduce health disparities? *American Behavioral Scientist*, 49(6), 775–793.
- Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1), 27–38.
- Friedman, A. L., Uhrig, J., Poehlman, J., Scales, M., & Hogben, M. (2014). Promoting sexual health equity in the United States: Implications from exploratory research with African-American adults. *Health Education Research*, 29(6), 993–1004. <https://doi.org/10.1093/her/cyu003>
- Frohlich, K. L., & Potvin, L. (2008). Transcending the known in public health practice: The inequality paradox: The population approach and vulnerable populations. *American Journal of Public Health*, 98(2), 216–221.
- GBD 2019 Diseases and Injuries Collaborators. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258), 1204–1222. [https://doi.org/10.1016/s0140-6736\(20\)30925-9](https://doi.org/10.1016/s0140-6736(20)30925-9)
- Goodstein, M. H., Lagon, E., Bell, T., Joyner, B. L., & Moon, R. Y. (2018). Stock photographs do not comply with infant safe sleep guidelines. *Clinical Pediatrics*, 57(4), 403–409.
- Hannon, L., & DeFina, R. (2016). Reliability concerns in measuring respondent skin tone by interviewer observation. *Public Opinion Quarterly*, 80(2), 534–541. <https://doi.org/10.1093/poq/nfw015>
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77–89. <https://doi.org/10.1080/19312450709336664>
- Institute of Medicine. (2002). *Speaking of health: Assessing health communication strategies for diverse populations*. National Academies Press. <https://doi.org/10.17226/10018>
- King, A. J. (2015). A content analysis of visual cancer information: Prevalence and use of photographs and illustrations in printed health materials. *Health Communication*, 30(7), 722–731. <https://doi.org/10.1080/10410236.2013.878778>
- Klonoff, E. A., & Landrine, H. (2000). Is skin color a marker for racial discrimination? Explaining the skin color-hypertension relationship. *Journal of Behavioral Medicine*, 23(4), 329–338. <https://doi.org/10.1023/a:1005580300128>
- Kreuter, M. W., Hovmand, P., Pfeiffer, D. J., Fairchild, M., Rath, S., Golla, B., & Casey, C. (2014). The “long tail” and public health: New thinking for addressing health disparities. *American Journal of Public Health*, 104(12), 2271–2278. <https://doi.org/10.2105/ajph.2014.302039>
- Kreuter, M. W., Lukwago, S. N., Bucholtz, R. D., Clark, E. M., & Sanders-Thompson, V. (2003). Achieving cultural appropriateness in health promotion programs: Targeted and tailored approaches. *Health Education Behavior*, 30(2), 133–146.
- Landrine, H., & Corral, I. (2014). Targeting cancer information to African Americans: The trouble with talking about disparities. *Journal of Health Communication*, 20(2), 196–203. <https://doi.org/10.1080/10810730.2014.920061>
- Loeb, S., Borno, H. T., Gomez, S., Ravenell, J., Myrie, A., Sanchez Nolasco, T., Byrne, N., Cole, R., Black, K., Stair, S., Macaluso, J. N., Walter, D., Siu, K., Samuels, C., Kazemi, A., Crocker, R., Sherman, R., Wilson, G., Griffith, D. M., & Langford, A. T. (2022). Representation in online prostate cancer content lacks racial and ethnic diversity: Implications for Black and Latinx men. *Journal of Urology*, 207(3), 559–564. <https://doi.org/10.1097/ju.00000000000002257>
- Louie, P., & Wilkes, R. (2018). Representations of race and skin tone in medical textbook imagery. *Social Science & Medicine*, 202, 38–42. <https://doi.org/10.1016/j.socscimed.2018.02.023>
- Massey, D. S., & Martin, J. A. (2003). *The New Immigrant Survey skin color scale*. <https://nis.princeton.edu/downloads/nis-skin-color-scale.pdf>
- Myers, T., Richardson, F., & Chung, J. E. (2019). Racial and ethnic makeup in hospital's social media and online platforms: Visual representation of diversity in images and videos of Washington, DC hospitals. *Journal of Health Communication*, 24(5), 482–491.
- Niederdeppe, J., Bigman, C. A., Gonzales, A. L., & Gollust, S. E. (2013). Communication about health disparities in the mass media. *Journal of Communication*, 63(1), 8–30. <https://doi.org/10.1111/jcom.12003>
- Parker, R., Larkin, T., & Cockburn, J. (2017). A visual analysis of gender bias in contemporary anatomy textbooks. *Social Science & Medicine*, 180, 106–113. <https://doi.org/10.1016/j.socscimed.2017.03.032>
- Resnicow, K., Baranowski, T., Ahluwalia, J. S., & Braithwaite, R. L. (1999). Cultural sensitivity in public health: Defined and demystified. *Ethnicity & Disease*, 9(1), 10–21.

Riffe, D., Lacy, S., & Fico, F. G. (2005). *Analyzing media messages: Using quantitative content analysis in research* (2nd ed.). Lawrence Erlbaum.

Skinner-Dorkenoo, A. L., Sarmal, A., Rogbeer, K. G., André, C. J., Patel, B., & Cha, L. (2022). Highlighting COVID-19 racial disparities can reduce support for safety precautions among White U.S. residents. *Social Science & Medicine*, 301, 114951. <https://doi.org/10.1016/j.socscimed.2022.114951>

Strait, A., Graf, J., Margaretten, M., Yazdany, J., & Goglin, S. (2021). Race, ethnicity and disparities in rheumatology educational materials. *Arthritis Care and Research*, 74(9), 1416–1420.

Strecher, V. J., McClure, J., Alexander, G., Chakraborty, B., Nair, V., Konkel, J., Greene, S., Couper, M., Carlier, C., Wiese, C., Little, R., Pomerleau, C., & Pomerleau, O. (2008). The role of engagement in a tailored web-based smoking cessation program: Randomized controlled trial. *Journal of Medical Internet Research*, 10(5), Article e36. <https://doi.org/10.2196/jmir.1002>

Webb Hooper, M., Antoni, M. H., Okuyemi, K., Dietz, N. A., & Resnicow, K. (2017). Randomized controlled trial of group-based culturally specific cognitive behavioral therapy among African American smokers. *Nicotine & Tobacco Research*, 19(3), 333–341. <https://doi.org/10.1093/ntr/ntw181>