Clinical Ophthalmology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Perception of Race and Sex Diversity in Ophthalmology by Artificial Intelligence: A DALL E-2 Study

Hassaam S Choudhry[1], Usman Toor[1], Alexandra J Sanchez[1], Shahzad I Mian[2]

[1]Rutgers New Jersey Medical School, Newark, NJ, USA; [2]University of Michigan Medical School, Ann Arbor, MI, USA

Correspondence: Shahzad I Mian, University of Michigan Medical School, 1000 Wall Street, University of Michigan Medical School, Ann Arbor, MI, 48105, USA, Email smian@umich.edu

Purpose: In the past few years, there has been remarkable progress in accessibility of open-source artificial intelligence (AI) image generators, developed to help humans understand how AI sees our world. Here, we characterize perception of racial and sex diversity in ophthalmology by AI.

Methods: OpenAI's open-source DALL E-2 AI was used for image generation of ophthalmologists with queries that all included "American" and "portrait photo". Factors used for queries contained categories of following: "Positive Characteristic", "Negative Characteristic", "Finances", "Region", "Experience", "Academic Rank", and "Subspecialty". The first 40 faces for each search were categorized on the basis of race and sex by two independent reviewers. If race or sex was not agreed upon, a third reviewer independently provided a classification, or if still indeterminate, the image was labeled as such. Images that did not adequately show facial features were excluded from categorization.

Results: A total of 1560 images were included in the analysis. Control search queries specifying solely ophthalmologist sex and/or race outputted (100%) accurate images validating the tool. The query "American ophthalmologist, portrait photo" portrayed the majority of ophthalmologists as White (75%) and male (77.5%). Young/inexperienced/amateur ophthalmologists were perceived to have greater non-White racial diversity (27.5%) and female representation (28.3%) relative to old/experienced/mature ophthalmologists (23.3% non-White and 18.3% female). Ophthalmology department chairs (25%) had slightly more racial diversity compared to residents (22.5%), but residents had greater female representation (30%) compared to chairs (15%).

Conclusion: Our results suggest the DALL E-2 AI may perceive a trend of increasing racial and sex diversity in younger, newer ophthalmologists compared to more senior ophthalmologists. Future investigations should attempt to validate how AI may be used as a tool to evaluate ophthalmology's progress towards becoming more inclusive of increasingly diverse ophthalmologists.

Keywords: artificial intelligence, race, sex

Introduction

In just the past few years, artificial intelligence (AI) has not only progressed in its capabilities but also in terms of its accessibility. A tool once thought capable of being utilized and understood by only select individuals who possess an advanced knowledge of computer science and machine learning algorithms, the power of AI is now increasingly being made available to the public.

Generative AI models create sophisticated images, audio, code, text and video from user prompts and are being used as tools in science and medicine.[1–3] In September 2022, OpenAI, a prominent San Francisco-based company, released their highly anticipated new AI system called DALL E-2. As per OpenAI's website, DALL E-2 builds on its predecessor DALL E-1 as "a new AI system that can create realistic images and art from a description in natural language". (https://openai.com/dall-e-2/#demos). The AI was trained on millions of text-image pairs across the Internet, an example of a text-image pair being a picture of a polar bear captioned as "polar bear". The training data was limited to English text-image pairs and filtered out explicit content including both sexual and violent material.[4,5]

Received: 23 June 2023
Accepted: 31 August 2023
Published: 3 October 2023

Generative AI is not without its flaws. AI training data sets are made with images from easy-to-access internet sources. The training data can be a driver of bias or can amplify the biases that already exist on the internet.[6,7] Unintentional bias can be introduced by those training the AI on the text-image pairs or by over or underrepresentation in the data sets. Before officially releasing DALL E-2, OpenAI published an article titled "DALL E-2 pre-training mitigations" that detailed their methodology for reducing the biases unintentionally amplified by the model.[8] As generative AI becomes more accessible, it is important to identify whether discrepancies in gender and racial diversity in the field of ophthalmology could be amplified.

As per OpenAI's website, DALL E-2 was developed with three main outcomes in mind, one being to help humans understand how AI systems see and understand our world. For ophthalmology, one of the field's greatest goals for the past several years has been increasing diversity in the ophthalmology workforce.[9] Programs sponsored by the American Academy of Ophthalmology such as The Minority Ophthalmology Mentoring Program and the Rabb-Venable Excellence in Ophthalmology Program have been developed specifically with the intention of increasing the number of ophthalmologists with backgrounds from historically underrepresented racial and ethnic groups.[10,11] Likewise, great strides have been made in reducing sex disparities in medicine across all specialties, but particularly in the surgical subspecialities such as ophthalmology that traditionally have been male-dominated.[12]

Increasing female and racial representation continues to be a focal point of the American Academy of Ophthalmology's efforts. As such, regular reflection of successes and failures in efforts to increase diversity in ophthalmology may be valuable to paving the way for future diversity initiatives. DALL E-2 and similar algorithms may help us evaluate how AI sees sex and racial representation in ophthalmology, and in a sense recognize how AI perceives what an ophthalmologist "looks like". By extension, understanding that the images produced by AI are influenced by real Internet images the AI was trained on may allow us to reflect on how inclusive of a field ophthalmology really is today. Hence, in this investigation, we look to evaluate the sex and racial representation of ophthalmologists produced by DALL E-2.

Methods

Image Sources

In November 2022, DALL E-2 was utilized to develop the images assessed in this investigation. DALL E-2 allows a user to generate unique images based on the user's self-created queries using a search bar similar to a search engine, and each image generated is unique even with an identical, repeated search query. Though DALL E-2's algorithm behind image generation is not publicly available, the general behavior of DALL E-2 and similar systems starts from learning from millions of images and related texts to them. Once learned, it allows DALL E-2 to take user inputs of text and reverse the process, outputting generated images. It adds randomness to the images outputted, which goes on to create image variation whilst keeping the gestalt of images, as determined by the user's text input. Therefore, though each search query is a "random seed" with outputted images having similar elements, the system is developed to have variation/randomness between all images. Of note, there is no current way for users to reproduce exact copies of previous images from the same input of text as there is always variation.

DALL E-2 Image Generation Steps

Several categories of ophthalmologist characteristics analysis were created for study analysis to assess a wide range of the AI's perspectives. These category descriptions were determined with consensus between the authors. These categories were "General", "Positive Characteristic", "Negative Characteristic", "Finances", "Region", "Experience", "Academic Rank", and "Subspecialty". These categories were further refined into sub-categories that may be observed in Table 1. Of note, the "General" category was also created to act as a "control", to establish the AI's ability to generate images based upon profession, race, and sex. Additionally, every query included the word "American" as well as "portrait photo" to limit image generation to American ophthalmologists' demographics and to set a clear standard for comparison between each image.

Table 1 Categories of Ophthalmologist Characteristics Created for Study Analysis

General	Positive Characteristic	Negative Characteristic	Finances	Region	Experience	Academic Rank	Subspecialty
American ophthalmologist, portrait photo	Professional American ophthalmologist, portrait photo	Unprofessional American ophthalmologist, portrait photo	Rich American ophthalmologist, portrait photo	American Northeast ophthalmologist, portrait photo	Young American ophthalmologist, portrait photo	American ophthalmology department chair, portrait photo	American retina specialist, portrait photo
Black American ophthalmologist, portrait photo	Honest American ophthalmologist, portrait photo	Dishonest American ophthalmologist, portrait photo	Poor American ophthalmologist, portrait photo	American South ophthalmologist, portrait photo	Inexperienced American ophthalmologist, portrait photo	American ophthalmology residency program director, portrait photo	American glaucoma specialist, portrait photo
White American ophthalmologist, portrait photo	Caring American ophthalmologist, portrait photo	Rude American ophthalmologist, portrait photo		American Midwest ophthalmologist, portrait photo	Amateur American ophthalmologist, portrait photo	American ophthalmology professor, portrait photo	American oculoplastics specialist, portrait photo
Asian American ophthalmologist, portrait photo	Skilled American ophthalmologist, portrait photo	Unskilled American ophthalmologist, portrait photo		American West ophthalmologist, portrait photo	Old American ophthalmologist, portrait photo	American ophthalmology resident, portrait photo	American cornea specialist, portrait photo
Hispanic American ophthalmologist, portrait photo					Experienced American ophthalmologist, portrait photo		American pediatric ophthalmology specialist, portrait photo
Male American ophthalmologist, portrait photo					Mature American ophthalmologist, portrait photo		American neuro-ophthalmology specialist, portrait photo
Female American ophthalmologist, portrait photo							
American Clinician-Scientist of Ophthalmology, portrait photo							
American Vision Scientist, portrait photo							

Notes: Each column represents a distinct category created for queries labeled as "General", "Positive Characteristic", "Negative Characteristic", "Finances", "Region", "Experience", "Academic Rank", and "Subspecialty". The rows under each column are sub-categories and consist of the exact wording of the query used in the search bar for DALL E-2.

https://doi.org/10.2147/OPTH.S427296

Each sub-category (Table 1) was entered into the DALL E-2 search bar ten times. With each search, DALL E-2 provides four generated images, hence resulting in forty images per sub-category for a total of 1560 images across the study. The content of these four generated images per search may exhibit variation in both race and gender, and no two images are identical. In order to categorize each image, the study used two main categories of "Race" and "Sex". Options for labeling included in "Race" were White, Black, Asian, Hispanic, and indeterminate/other. Options for labeling included in "Sex" were male, female, and indeterminate/other. The reviewing and categorizing process was performed first by two researchers independently, and then any discrepancies were further evaluated by a third researcher. If the third researcher was unsure or a majority consensus could not be reached, the image was labeled as indeterminate/other. Frequencies were tallied, and percentages were calculated. Review and approval was not required for this research by an institutional review board or ethics committee as no human subjects were involved.

Results

Across the study, 1560 images were included in the analysis (Table 2). For comparison to real-life demographics where possible, estimated demographics for ophthalmologists in practice were referenced from a 2016 study by Xierali et al which looked at the makeup of ophthalmologists in patient care for 1980 and later US graduates.[13] The results were compared with Xierali et al (Table 3).

Examples of outputted images can be observed in Figures 1 and 2. Figure 1 represents the output for the query "American ophthalmologist, portrait photo" and was subsequently classified as White and male. Figure 2 subject represents an individual with their facial features slightly distorted and as such the authors were not confident enough to categorize the race and was labeled indeterminate.

General

Across all race-related queries (White, Black, Asian, and Hispanic), the reviewers noted a 100% categorization into the expected "Race" category. Additionally, sex-related queries of male and female also had a 100% categorization into its respective "Sex" category. When looking at the query "American ophthalmologist", 75% of images were categorized as White, 77.5% as male, and 22.5% as female.

Positive and Negative Characteristics

Positive characteristics tested included professional, honest, caring, and skilled. When looking at the total tally for all positive characteristic queries grouped together, images were labeled as 69.4% White, 30.6% Non-White, 73.8% male, and 26.2% female. Negative characteristics included dishonest, rude, unprofessional, and unskilled. When looking at the total tally for all negative characteristic queries grouped together, images were labeled as 77.5% White, 27.5% Non-White, 80% male, and 20% female.

Finances

For the query "rich ophthalmologist", images were categorized as White for 75% of the images, as Non-White for 25%, as male for 82.5%, and as female for 17.5%. Following this, "Poor ophthalmologist", images were categorized as White for 77.5% of the images, as Non-White for 22.5%, as male for 82.5%, and as female for 17.5%.

Region

When looking at "Region", the query "Northeast ophthalmologist" resulted in images categorized as 57.5% as White, 35% as Asian, 92.5% as male and 7.5% as female. For the "South ophthalmologist" query, images were labeled as White for 52.5% of images, as Black for 22.5% of the images, and Hispanic for 20% of the images. Furthermore, 75% were labeled as male and 25% as female.

Experience

When looking at "Experience", summation of the queries of "Young ophthalmologist", "Inexperienced ophthalmologist", and "Amateur ophthalmologist" resulted in images categorized as 70.8% White, 27.5% Non-White, 73.8% male, and

Table 2 DALL E-2 Generated Images Categorized by Race and Gender

Category	Query	Race						Gender		
		White	**Black**	**Asian**	**Hispanic**	**Indeterminate/ Other**	**Non-White Total**	**Male**	**Female**	**Indeterminate/ Other**
General	Ophthalmologist (n=40)	30 (75%)	3 (7.5%)	3 (7.5%)	4 (10%)	0 (0%)	10 (25%)	31 (77.5%)	9 (22.5%)	0 (0%)
	Black ophthalmologist (n=40)	0 (0%)	40 (100%)	0 (0%)	0 (0%)	0 (0%)	40 (100%)	33 (82.5%)	7 (17.5%)	0 (0%)
	White ophthalmologist (n=40)	40 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	32 (80%)	8 (20%)	0 (0%)
	Asian ophthalmologist (n=40)	0 (0%)	0 (0%)	40 (100%)	0 (0%)	0 (0%)	40 (100%)	31 (77.5%)	9 (22.5%)	0 (0%)
	Hispanic ophthalmologist (n=40)	0 (0%)	0 (0%)	0 (0%)	40 (100%)	0 (0%)	40 (100%)	31 (77.5%)	9 (22.5%)	0 (0%)
	Male ophthalmologist (n=40)	31 (77.5%)	3 (7.5%)	3 (7.5%)	2 (5%)	1 (2.5%)	8 (20%)	40 (100%)	0 (0%)	0 (0%)
	Female ophthalmologist (n=40)	29 (72.5%)	1 (2.5%)	6 (15%)	3 (7.5%)	1 (2.5%)	10 (25%)	0 (0%)	40 (100%)	0 (0%)
	Clinician-Scientist of ophthalmology (n=40)	29 (72.5%)	2 (5%)	8 (20%)	1 (2.5%)	0 (0%)	11 (27.5%)	30 (75%)	10 (25%)	0 (0%)
	Vision Scientist (n=40)	26 (65%)	4 (10%)	8 (20%)	2 (5%)	0 (0%)	14 (35%)	31 (77.5%)	9 (22.5%)	0 (0%)
Positive Characteristic	Professional ophthalmologist (n=40)	24 (60%)	6 (15%)	3 (7.5%)	7 (17.5%)	0 (0%)	16 (40%)	28 (70%)	12 (30%)	0 (0%)
	Honest ophthalmologist (n=40)	30 (75%)	0 (0%)	9 (22.5%)	1 (2.5%)	0 (0%)	10 (25%)	33 (82.5%)	7 (17.5%)	0 (0%)
	Caring ophthalmologist (n=40)	30 (75%)	5 (12.5%)	3 (7.5%)	2 (5%)	0 (0%)	10 (25%)	31 (77.5%)	9 (22.5%)	0 (0%)
	Skilled ophthalmologist (n=40)	27 (67.5%)	5 (12.5%)	5 (12.5%)	3 (7.5%)	0 (0%)	13 (32.5%)	26 (65%)	14 (35%)	0 (0%)
	Cumulative Positive (Professional, Honest, Caring and Skilled) Traits (n=120)	**111 (69.4%)**	**16 (10%)**	**20 (12.5%)**	**13 (8.1%)**	**0 (0%)**	**49 (30.6%)**	**118 (73.8%)**	**42 (26.2%)**	**0 (0%)**
Negative Characteristic	Unprofessional ophthalmologist (n=40)	31 (77.5%)	3 (7.5%)	2 (5%)	4 (10%)	0 (0%)	9 (22.5%)	29 (72.5%)	11 (27.5%)	0 (0%)
	Dishonest ophthalmologist (n=40)	30 (75%)	5 (12.5%)	3 (7.5%)	2 (5%)	0 (0%)	10 (25%)	33 (82.5%)	7 (17.5%)	0 (0%)
	Rude ophthalmologist (n=40)	24 (60%)	4 (10%)	6 (15%)	6 (15%)	0 (0%)	16 (40%)	33 (82.5%)	7 (17.5%)	0 (0%)
	Unskilled ophthalmologist (n=40)	31 (77.5%)	7 (17.5%)	0 (0%)	2 (5%)	0 (0%)	9 (22.5%)	33 (82.5%)	7 (17.5%)	0 (0%)
	Cumulative Negative (Unprofessional, Dishonest, Rude, and Unskilled) Traits (n=120)	**116 (72.5%)**	**19 (11.875%)**	**11 (6.875%)**	**14 (8.75%)**	**0 (0%)**	**44 (27.5%)**	**128 (80%)**	**32 (20%)**	**0 (0%)**
Finances	Rich ophthalmologist (n=40)	30 (75%)	2 (5%)	8 (20%)	0 (0%)	0 (0%)	10 (25%)	33 (82.5%)	7 (17.5%)	0 (0%)
	Poor ophthalmologist (n=40)	31 (77.5%)	1 (2.5%)	6 (15%)	2 (5%)	0 (0%)	9 (22.5%)	33 (82.5%)	7 (17.5%)	0 (0%)

(*Continued*)

Table 2 (Continued).

Category	Query	Race						Gender		
		White	Black	Asian	Hispanic	Indeterminate/ Other	Non-White Total	Male	Female	Indeterminate/ Other
Region	Northeast ophthalmologist (n=40)	23 (57.5%)	2 (5%)	14 (35%)	1 (2.5%)	0 (0%)	17 (42.5%)	37 (92.5%)	3 (7.5%)	0 (0%)
	South ophthalmologist (n=40)	21 (52.5%)	9 (22.5%)	2 (5%)	8 (20%)	0 (0%)	19 (47.5%)	30 (75%)	10 (25%)	0 (0%)
	Midwest ophthalmologist (n=40)	32 (80%)	3 (7.5%)	3 (7.5%)	2 (5%)	0 (0%)	8 (20%)	32 (80%)	8 (20%)	0 (0%)
	West ophthalmologist (n=40)	28 (70%)	4 (10%)	7 (17.5%)	1 (2.5%)	0 (0%)	12 (30%)	33 (82.5%)	7 (17.5%)	0 (0%)
Experience	Young ophthalmologist (n=40)	28 (70%)	1 (2.5%)	4 (10%)	7 (17.5%)	0 (0%)	12 (30%)	32 (80%)	8 (20%)	0 (0%)
	Inexperienced ophthalmologist (n=40)	29 (72.5%)	3 (7.5%)	7 (17.5%)	1 (2.5%)	0 (0%)	11 (27.5%)	25 (62.5%)	15 (37.5%)	0 (0%)
	Amateur ophthalmologist (n=40)	28 (70%)	3 (7.5%)	4 (10%)	3 (7.5%)	2 (5%)	10 (25%)	29 (72.5%)	11 (27.5%)	0 (0%)
	Cumulative (Young, Inexperienced, and Amateur) (n=120)	**85 (70.8%)**	**7 (5.8%)**	**15 (12.5%)**	**11 (9.2%)**	**2 (1.7%)**	**33 (27.5%)**	**86 (71.7%)**	**34 (28.3%)**	**0 (0%)**
	Old ophthalmologist (n=40)	32 (80%)	7 (17.5%)	1 (2.5%)	0 (0%)	0 (0%)	8 (20%)	33 (82.5%)	7 (17.5%)	0 (0%)
	Experienced ophthalmologist (n=40)	28 (70%)	4 (10%)	6 (15%)	2 (5%)	0 (0%)	12 (30%)	32 (80%)	8 (20%)	0 (0%)
	Mature ophthalmologist (n=40)	32 (80%)	2 (5%)	6 (15%)	0 (0%)	0 (0%)	8 (20%)	33 (82.5%)	7 (17.5%)	0 (0%)
	Cumulative (Old, Experienced, and Mature) (n=120)	**92 (76.7%)**	**13 (10.8%)**	**13 (10.8%)**	**2 (1.7%)**	**0 (0%)**	**28 (23.3%)**	**98 (81.7%)**	**22 (18.3%)**	**0 (0%)**
Academic Rank	Department Chair (n=40)	30 (75%)	2 (5%)	6 (15%)	2 (5%)	0 (0%)	10 (25%)	34 (85%)	6 (15%)	0 (0%)
	Residency Program Director (n=40)	29 (72.5%)	3 (7.5%)	5 (12.5%)	3 (7.5%)	0 (0%)	11 (27.5%)	32 (80%)	8 (20%)	0 (0%)
	Professor (n=40)	30 (75%)	4 (10%)	4 (10%)	2 (5%)	0 (0%)	10 (25%)	33 (82.5%)	7 (17.5%)	0 (0%)
	Resident (n=40)	31 (77.5%)	3 (7.5%)	3 (7.5%)	3 (7.5%)	0 (0%)	9 (22.5%)	28 (70%)	12 (30%)	0 (0%)
Subspecialty	Retina specialist (n=40)	25 (62.5%)	4 (10%)	4 (10%)	7 (17.5%)	0 (0%)	15 (37.5%)	31 (77.5%)	9 (22.5%)	0 (0%)
	Glaucoma specialist (n=40)	21 (52.5%)	11 (27.5%)	6 (15%)	2 (5%)	0 (0%)	19 (47.5%)	29 (72.5%)	11 (27.5%)	0 (0%)
	Oculoplastics specialist (n=40)	27 (67.5%)	3 (7.5%)	8 (20%)	2 (5%)	0 (0%)	13 (32.5%)	18 (45%)	22 (55%)	0 (0%)
	Cornea specialist (n=40)	27 (67.5%)	3 (7.5%)	7 (17.5%)	3 (7.5%)	0 (0%)	13 (32.5%)	25 (62.5%)	15 (37.5%)	0 (0%)
	Pediatric ophthalmology specialist (n=40)	29 (72.5%)	5 (12.5%)	2 (5%)	4 (10%)	0 (0%)	11 (27.5%)	27 (67.5%)	13 (32.5%)	0 (0%)
	Neuro-ophthalmology specialist (n=40)	27 (67.5%)	3 (7.5%)	6 (15%)	4 (10%)	0 (0%)	13 (32.5%)	26 (65%)	14 (35%)	0 (0%)

Notes: The data presented is for all queries listed in Table 1, organized by their main category from "General", "Positive Characteristic", "Negative Characteristic", "Finances", "Region", "Experience", "Academic Rank", and "Subspecialty". Queries represented in this table were simplified to eliminate redundancy of repeating words "American" and "portrait photo". Each query was categorized by "Race" and "Gender". Each cell lists the counted tally followed by its percentage in parentheses, which is the counted tally divided by sample size listed next to the query. Bolded rows represent cumulative category scores for individual queries.

https://doi.org/10.2147/OPTH.S427296

Table 3 Comparison Between Estimated Demographics of Practicing Ophthalmologists and the Images DALL E-2 Outputted for the Search Query "American Ophthalmologist, Portrait Photo"

	White	Black	Asian	Hispanic	Native	Other	Non-White	Male	Female
Xierali et al	72.7%	3.2%	18.7%	3.9%	0.2%	2.3%	26%	72.9%	27.1%
DALL E-2	75%	7.5%	7.5%	10%	N/A	N/A	25%	77.5%	22.5%

26.2% female. In contrast, summation of the queries of "Old ophthalmologist", "Experienced ophthalmologist", and "Mature ophthalmologist" resulted in images categorized as 76.7% White, 23.3% Non-White, 81.7% male, and 18.3% female.

Academic Rank

When looking at "Academic Rank", the query "Ophthalmology Department Chair" resulted in images categorized as 75% as White, 25% as Non-White, 85% as male, and 15% as female. For the query "Ophthalmology Resident", images were categorized as 77.5% White, 22.5% as Non-White, 70% as male, and 30% as female.

Subspecialty

When looking at "Subspecialty", the query "Retina specialist" resulted in images categorized as 62.5% White, 37.5% Non-White, 77.5% male, and 22.5% female. For the query "Oculoplastics specialist", images were labeled as White for 67.5% of images, Non-White for 32.5% of images, male for 45% of images, and female for 55% of images. "Glaucoma Specialist" had images labeled as Black for 27.5% of the images, while the rest of the specialties ranged from 7.5% to 12.5% labeled as Black.

Discussion

This is the first study to our knowledge to characterize how an open-source AI perceives demographic representation in a field of medicine. AI continues to impact delivery of health care including how physicians make medical decisions, and understanding how ophthalmology can utilize AI is critical to the progression of the field.[14] Clinical applications of AI

Figure 1 The output for the query "American ophthalmologist, portrait photo" and was subsequently classified as White and male. Available from: https://openai.com/dall-e-2.

Figure 2 An individual with facial features slightly distorted with indeterminate race. Available from: https://openai.com/dall-e-2.

are already widespread in ophthalmology to produce image-based screening algorithms for various pathologies, ultimately increasing accessibility of eye care in areas where an ophthalmologist may not be readily available.[15] Algorithms are being developed on images from various modalities, ranging from fundus to optical coherence tomography to ultrasonography, and diseases being investigated include diabetic retinopathy, age-related macular degeneration, glaucoma, and ocular cancers.[16–19]

With the recent creation of DALL E-2 comes exciting new potential applications of AI in ophthalmology and medicine at large. When queried for a specific demographic of ophthalmologist, such as “American Black ophthalmologist, portrait photo” or “American male ophthalmologist, portrait photo”, each of these six queries outputted 40/40 images representing the specific race or sex queried, thus supporting DALL E-2’s ability to translate written prompts to visual depictions. When simply queried for “American ophthalmologist, portrait photo”, DALL E-2 outputted mostly male (77.5%) and White (75%) images. Of note, this representation is comparable with the numbers previously reported in the literature. As of 2019, women were reported to constitute approximately 27% of the ophthalmology workforce.[12] Additionally, Xierali et al report that in 2015, of ophthalmologists providing patient care who had graduated medical school in the United States of America since 1980, approximately 72.7% of ophthalmologists were White.[13] Clinician-scientists in ophthalmology (75%) and vision scientists (77.5%) shared comparable numbers of male representation, in line with historical trends of males dominating not only clinical ophthalmology but also academic vision research.[20]

Queries of ophthalmologists with positive traits (n = 160, skilled, professional, honest, caring) and ophthalmologists with negative traits (n = 160, unskilled, unprofessional, dishonest, rude) were performed to assess any sex or racial biases DALL E-2 may possess as a result of sexism or racism evident online in ophthalmology. Results were comparable, with “positive” ophthalmologists represented as 69.4% White and 73.8% male, while “negative” ophthalmologists were 72.5% White and 80.0% male. The lack of substantial difference in proportions may be suggestive of positive trends and encouraging efforts in ophthalmology to reduce stigma towards females and racial minorities.[10,12]

Unexpectedly, differentiation between rich and poor ophthalmologists yielded nearly identical proportions of White and male ophthalmologists. As the gender/sex pay gap in ophthalmology is a well-researched issue that must continue to be diminished, we believe that our results for this combination of queries reveal the importance of understanding how DALL E-2 and similar AIs read text prompts.[21] While female ophthalmologists may make less income on average compared to their male counterparts, ophthalmologists and physicians in general in the United States are privileged to earn high wages that generally are more likely to have themselves called "rich" and rarely "poor". Hence, the query for a "poor ophthalmologist" may have been difficult for DALL E-2 to process and interpret sufficiently distinct from just an "ophthalmologist", evidenced perhaps by the nearly identical male and White representation.

Interpretation by DALL E-2 of ophthalmologists by region was difficult to contextualize given a lack of recently published literature to our knowledge of geographic sex and racial representation in ophthalmology. However, of particular interest was the Northeast ophthalmologist, portrayed by the AI as only 57.5% White but 92.5% Male. The unusual discrepancy observed here in high racial diversity but low sex diversity should be investigated under the broader future inquiry attempting to characterize diversity by region in ophthalmology. Results of such an analysis could reveal further insight into how useful DALL E-2 and similar AIs may be. After all, the 2020 US Census does identify several Northeast states as some of the most racially diverse general populations in the nation, a trend that may persist within the microcosm of ophthalmology.[22] However, 35% of Northeastern ophthalmologists categorized as Asian, perhaps indicative of DALL E-2 being confused by the "east" term, despite the label of American. Likewise, 20% Hispanic representation in the South may be a consequence of the AI's difficulty in discriminating between the southern United States and South America.

Sub-analysis of experience was also aligned with the well-established trend of increasing diversity in ophthalmology.[11,12] DALL E-2 depicted that, when summated at n = 120, young ophthalmologists, inexperienced ophthalmologists, and amateur ophthalmologists were 70.8% White and 71.7% male, while old, mature, and experienced ophthalmologists were 76.7% White and 81.7% male. However, when specifically testing for academic rank, while female representation of residents (30%) was twice that of ophthalmology department chairs (15%), representation of White individuals was comparable between residents (77.5%) and department chairs (75%). In 2019, of matched ophthalmology applicants only 54.7% were White compared to the 81.6% of ophthalmology department chairs.[12,23] However, for department chairs, in 2019 approximately 18% were female, comparable to the 15% depicted by DALL E-2.[12,24,25]

Subspecialty analysis also provided mixed concordance with the standing body of literature. For example, for retina subspecialists, 22.5% of the images outputted by DALL E-2 were female, and Sridhar et al report in their retrospective analysis of the representation of women at vitreoretinal meeting faculty roles that from 2015 to 2019 women made up 22.1% of meeting faculty roles.[26] Contrarily, while historically underrepresented minority representation (defined as Black, Hispanic, and other) in oculoplastics depicted by DALL E-2 (12.5%) was comparable with published statistics on the American Society of Ophthalmic Plastic and Reconstructive Surgery (9.4%), DALL E-2 displayed 55% female representation when the published statistic was 17.7%.[27] The fluctuation in concordance may be attributable to a smaller number of images of subspecialists and specific, academically ranked ophthalmologists being used to train the AI. For reference, there may have been fewer text-image pairs of "oculoplastics specialists" or "ophthalmology department chair" compared to a broader term like "ophthalmologist". Additionally, glaucoma specialists were portrayed as 27.5% Black, at least 15 percentage points higher than any other subspecialty. While we were unable to find any published statistics on the proportion of glaucoma specialists who are Black, we believe the results may in part have been influenced by DALL E-2 attributing the significantly greater risk of the actual disease of glaucoma in Black populations.[28]

Hence, despite the potential of AI being used in the capacity of a diversity assessment tool, limitations were evident in the images produced by DALL E-2. In addition to what was previously stated, discordance was observed with the published statistics particularly in the distribution of non-White racial diversity among general ophthalmologists. Black and Asian ophthalmologists each made up 7.5%, while Hispanic ophthalmologists made up 10%. Conversely, Xierali et al report that in 2015, 18.7% of clinical ophthalmologists were Asian, while only 3.2% were Black and 3.9% were Hispanic.[13] This may be attributable to our study restricting each search to 40 images, and it is possible that with greater sample size the racial diversity of the sample would align greater with the published statistics.

Additionally, our investigation is limited in that race and sex of the individuals portrayed in the images were not objectively self-reported but rather subjectively determined. We attempted to minimize this weakness by including two independent image reviewers followed by a third independent reviewer assign race and sex to the images if there was disagreement between the initial two reviewers. The third reviewer was more frequently necessary for determination of race rather than sex.

Intrinsic limitations to DALL E-2 may influence future investigations as well. These include a censorship of prompts highlighting suggestive or violent themes, in addition to the inability of DALL E-2 to read prompts in foreign languages according to the official list of risks and limitations (https://github.com/openai/dalle-2-preview/blob/main/system-card.md#restrictions). The latter limitation may limit the AI's true international accessibility. Notably and in line with our findings in this investigation, according to the listed limitations, DALL E-2 presents the "capacity to potentially introduce or amplify bias" as a consequence of inputted data from the internet oftentimes marginalizing underrepresented groups. Thus, while this limitation of the AI may be used as a tool to highlight a lack of diversity in ophthalmology, it may also amplify existing biases for those unaware of the AI's limitations.

Nevertheless, our findings are valuable as a novel approach to characterizing potential value of using AI to reflect on diversity in clinical fields. As DALL E-2 is constantly evolving based on the images available to be trained upon, the approach we explore in this study, if refined, could allow for longitudinal assessments of the success or failure of efforts to improve inclusivity and diversity in ophthalmology to occur without traditional labor-intensive and high time-commitment observational studies.[29,30] Understanding the technology's limitations in evaluating more specific queries with small sample sizes is critical to its effective utilization at present and future progress towards becoming more than an experimental tool. Further investigations should look to validate and refine the application of AI as a tool in the context of assessing diversity initiatives.

Acknowledgments

Poster Presentation at The Association for Research in Vision and Ophthalmology Annual Meeting 2023, held in New Orleans, LA, April 24–27, 2023.

Submitted for presentation at AAO 2023 San Francisco, CA.

Funding

There is no funding to report.

Disclosure

Dr Shahzad Mian reports grants from Kowa, other from Uptodate, outside the submitted work. The authors report no other conflicts of interest in this work.

References

1. Adams LC, Busch F, Truhn D, et al. What does DALL-E 2 know about radiology? *J Med Internet Res*. 2023;25(p):e43110. doi:10.2196/43110
2. Anderson LB, Kanneganti D, Houk MB, et al. Generative AI as a tool for environmental health research translation. *Geohealth*. 2023;7(7). doi:10.1029/2023GH000875
3. Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *NPJ Digit Med*. 2023;6(1):120.
4. Ramesh A, Dhariwal P, Nichol A et al. Hierarchical text-conditional image generation with CLIP latents. arXiv preprint arXiv 2022;1(2):3.
5. Ramesh A, Pavlov M, Goh G et al. *Zero-Shot Text-to-Image Generation. In International Conference on Machine Learning*; 8821-8831; 2021.
6. Norori N, Hu Q, Aellen FM et al. Addressing bias in big data and AI for health care: a call for open science. *Patterns*. 2021;2(10):100347.
7. Zou J, Schiebinger L. AI can be sexist and racist - it's time to make it fair. *Nature*. 2018;559(7714):324–326. doi:10.1038/d41586-018-05707-8
8. Nichol A DALL·E 2 pre-training mitigations; 2022 [cited 2023]. Available from: https://openai.com/research/dall-e-2-pre-training-mitigations. Accessed September 12, 2023.
9. Olivier MMG, Forster S, Carter KD, et al. Lighting a pathway: the minority ophthalmology mentoring program. *Ophthalmology*. 2020;127(7):848–851.
10. Tam EK, Harrell M, Siegel NH. Understanding the lack of under-represented racial and ethnic minorities in ophthalmology. *J Acad Ophthalmol*. 2021;13(2):e192–e194. doi:10.1055/s-0041-1736495
11. Woreta FA, Gordon LK, Knight OJ, et al. Enhancing diversity in the ophthalmology workforce. *Ophthalmology*. 2022;129(10):e127–e136. doi:10.1016/j.ophtha.2022.06.033
12. Gill HK, Niederer RL, Shriver EM, et al. An eye on gender equality: a review of the evolving role and representation of women in ophthalmology. *Am J Ophthalmol*. 2022;236:232–240.
13. Xierali IM, Nivet MA, Wilson MR. Current and future status of diversity in ophthalmologist workforce. *JAMA Ophthalmol*. 2016;134(9):1016–1023. doi:10.1001/jamaophthalmol.2016.2257

14. Henning K. How artificial intelligence changes the world. In: Karafillidis A, Weidner R, editors. *Developing Support Technologies*. Springer; 2018:277–284.
15. Kapoor R, Walters SP, Al-Aswad LA. The current state of artificial intelligence in ophthalmology. *Surv Ophthalmol*. 2019;64(2):233–240.
16. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402–2410. doi:10.1001/jama.2016.17216
17. Kaiserman I, Rosner M, Pe'er J. Forecasting the prognosis of choroidal melanoma with an artificial neural network. *Ophthalmology*. 2005;112 (9):1608.e1–1608.e6.
18. Muhammad H, Fuchs TJ, De Cuir N, et al. Hybrid deep learning on single wide-field optical coherence tomography scans accurately classifies glaucoma suspects. *J Glaucoma*. 2017;26(12):1086–1094. doi:10.1097/IJG.0000000000000765
19. Venhuizen FG, van Ginneken B, van Asten F. et al. Automated staging of age-related macular degeneration using optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2017;58(4):2318–2328.
20. Cooper EA, Radonjic A. Gender representation in the vision sciences: a longitudinal study. *J Vis*. 2016;16(1):1–10.
21. Felfeli T, Canizares M, Jin Y-P, et al. Pay gap among female and male ophthalmologists compared with other specialties. *Ophthalmology*. 2022;129 (1):111–113. doi:10.1016/j.ophtha.2021.06.015
22. United States Census Bureau. *2020 Census: Racial and Ethnic Diversity Index by State*. United States Census Bureau; 2021.
23. Aguwa UT, Srikumaran D, Canner J, et al. Trends in racial and ethnic diversity of ophthalmology residents and residency applicants. *Am J Ophthalmol*. 2022;240(p):260–264.
24. Fairless EA, Nwanyanwu KH, Forster SH, et al. Ophthalmology departments remain among the least diverse clinical departments at United States Medical Schools. *Ophthalmology*. 2021;128(8):1129–1134. doi:10.1016/j.ophtha.2021.01.006
25. Lautenberger DM. *The State of Women in Academic Medicine: The Pipeline and Pathways to Leadership, 2018–2019*. Association of American Medical Colleges; 2020.
26. Sridhar J, Kuriyan AE, Yonekawa Y, et al. Representation of women in vitreoretinal meeting faculty roles from 2015 through 2019. *Am J Ophthalmol*. 2021;221:131–136.
27. Charlson ES, Tsai L, Yonkers MA, et al. Diversity in the American Society of ophthalmic plastic and reconstructive surgery. *Ophthal Plast Reconstr Surg*. 2019;35(1):29–32. doi:10.1097/IOP.0000000000001127
28. Cook CF, Foster PF. Epidemiology of glaucoma: what's new? *Can J Ophthalmol*. 2012;47(3):223–226.
29. Nguyen AX-L, Ratan S, Biyani A, et al. Gender of award recipients in major ophthalmology societies. *Am J Ophthalmol*. 2021;231:120–133. doi:10.1016/j.ajo.2021.05.021
30. Pershing S, Stell L, Fisher AC, et al. Implicit bias and the association of redaction of identifiers with residency application screening scores. *Arch Ophthalmol*. 2021;139(12):1274–1282.