"Making Waves"
25th Annual Pain Medicine Meeting
November 5-7, 2026 | Tampa, Florida



School of Medicine and Public Health
UNIVERSITY OF WISCONSIN-MADISON

Generative AI Representation of the Anesthesiology Workforce: A Comparison with National Workforce Demographics

UW SMPH - Department of Anesthesiology
Jesus Lopez Ortega, BA, Balkaran Sing, BS, Rachel E. Corbitt, MD, Doris Pierson, MD, Lee W. Conway, BS, Yazan Hammad, BS, Sara Eagle, Debo Paul, BS, Cora Schroeder, Kristopher Schroeder, MD

Introduction

Generative AI models learn associations from datasets that may reflect societal biases, raising concerns about how physicians are represented. If AI-generated images differ from the actual physician workforce, they may reproduce or amplify demographic imbalances and stereotypes. This study evaluated race, ethnicity, gender, and age in AI-generated images across five anesthesiology subspecialties and compared them with national AAMC and ACGME workforce demographics.

Methods

We generated 75 images across five anesthesiology subspecialties using three AI models, ChatGPT, Grok, and Gemini. Nine independent raters classified 732 depicted physicians by gender, race/ethnicity, and age. Demographic proportions were calculated for each image and averaged by subspecialty. Inter-rater reliability was assessed using the Intraclass Correlation Coefficient, and AI-generated proportions were compared with national AAMC/ACGME workforce benchmarks using one-sample t-tests.

Results

AI-generated demographic proportions differed significantly from real-world ACGME and AAMC workforce benchmarks across nearly all subspecialties, with the most consistent pattern being overrepresentation of White and male physicians. One-sample t-tests demonstrated statistically significant deviations in multiple demographic comparisons.

Other significant findings:

- Obstetric Anesthesiology: Women were underrepresented in AI-generated images (35.0% AI vs 64.4% workforce; $p < 0.001$).
- Critical Care Anesthesiology: Women were underrepresented (24.8% AI vs 46.3% workforce; $p < 0.001$).
- Pediatric Anesthesiology: Women were underrepresented (31.0% AI vs 51.0% workforce; $p < 0.001$), while White physicians were overrepresented (82.2% AI vs 56.9% workforce; $p < 0.001$).
- Pain Medicine: White physicians were overrepresented (80.6% AI vs 41.2% workforce; $p < 0.001$).
- Cardiac Anesthesiology: White physicians were overrepresented (81.0% AI vs 48.9% workforce; $p < 0.001$).

Results

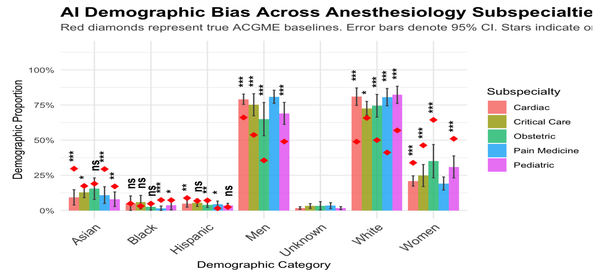


Figure 1: Colored bars represent mean AI-generated demographic proportions by subspecialty; error bars show 95% confidence intervals, red diamonds indicate real-world workforce benchmarks, and asterisks denote statistical significance.

Discussion

AI-generated images differed substantially from real-world anesthesiology workforce demographics, consistently overrepresenting White and male physicians. Notably, these patterns persisted even in subspecialties that are female-dominated or nearly gender-balanced in the actual workforce. These discrepancies may reflect historical demographic imbalances and stereotypes learned from training data that generative AI models reproduce or amplify. As anesthesiology increasingly explores emerging AI technologies, recognizing and addressing these biases will be important to ensure that future applications accurately and equitably represent the physicians and communities the specialty serves.