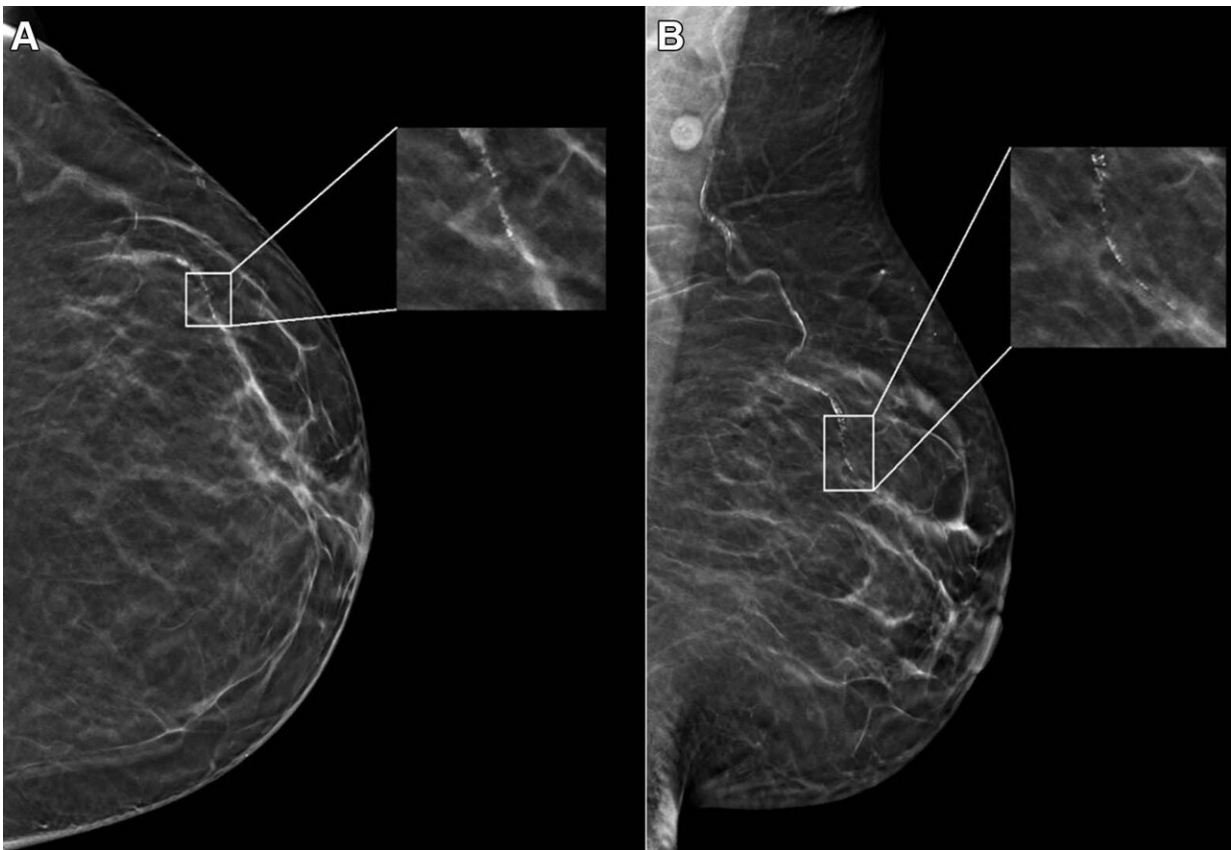


# Age, race impact AI performance on digital mammograms, study finds

May 21 2024



Example mammogram assigned a false-positive case score of 96 in a 59-year-old Black patient with scattered fibroglandular breast density. (A) Left craniocaudal and (B) mediolateral oblique views demonstrate vascular calcifications in the upper outer quadrant at middle depth (box) that were singularly identified by the artificial intelligence algorithm as a suspicious finding and assigned an individual lesion score of 90. This resulted in an overall case score assigned to the mammogram of 96. Credit: Radiological Society of North America (RSNA)

In a study of nearly 5,000 screening mammograms interpreted by an FDA-approved AI algorithm, patient characteristics such as race and age influenced false positive results. The study's results were published in *Radiology*.

"AI has become a resource for radiologists to improve their efficiency and accuracy in reading [screening mammograms](#) while mitigating reader burnout," said Derek L. Nguyen, M.D., assistant professor at Duke University in Durham, North Carolina. "However, the impact of patient characteristics on AI performance has not been well studied."

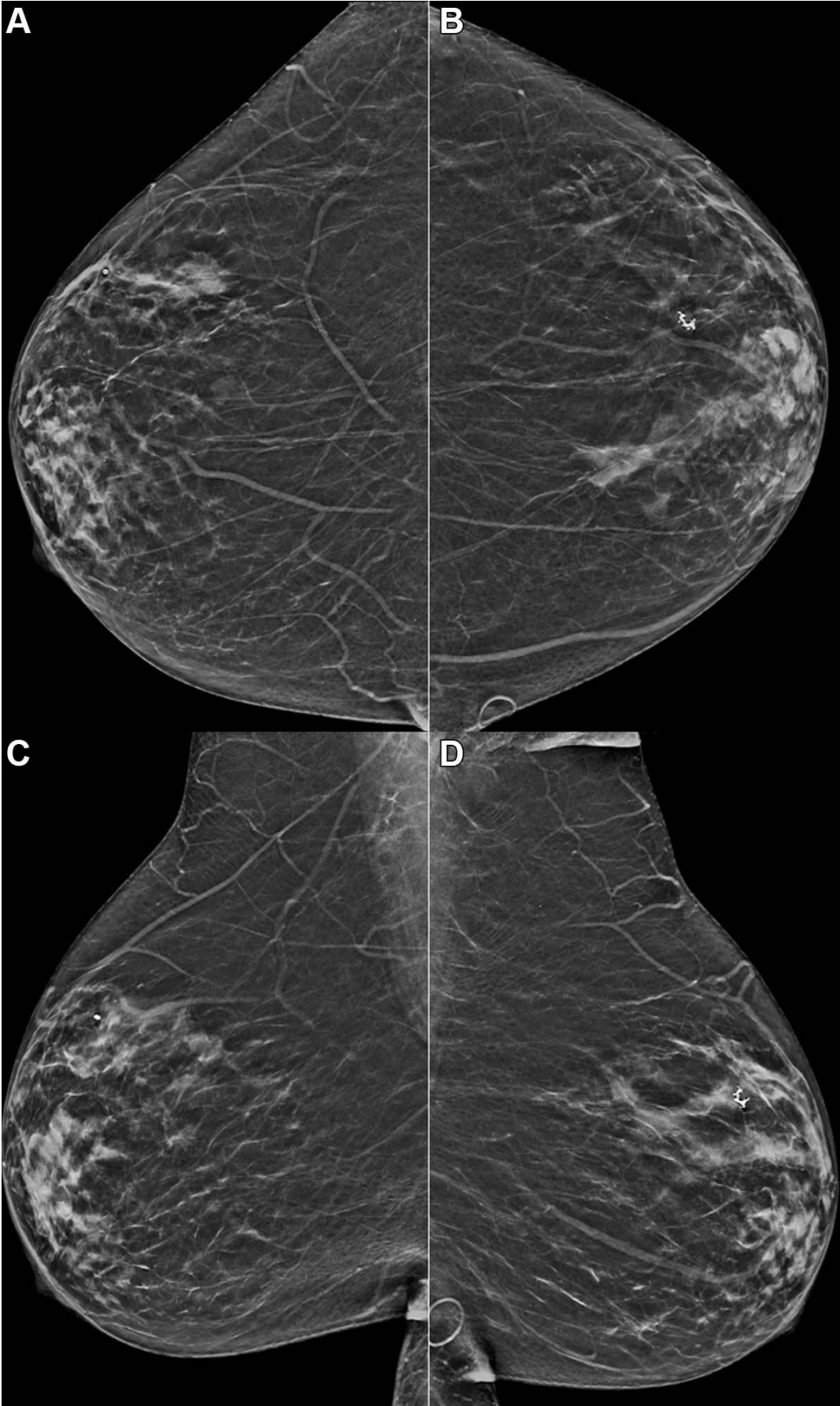
Dr. Nguyen said while preliminary data suggests that AI algorithms applied to screening mammography exams may improve radiologists' diagnostic performance for [breast cancer detection](#) and reduce interpretation time, there are some aspects of AI to be aware of.

"There are few demographically diverse databases for AI algorithm training, and the FDA does not require diverse datasets for validation," he said. "Because of the differences among patient populations, it's important to investigate whether AI software can accommodate and perform at the same level for different patient ages, races and ethnicities."

In the [retrospective study](#), researchers identified patients with negative (no evidence of cancer) digital breast tomosynthesis screening examinations performed at Duke University Medical Center between 2016 and 2019. All patients were followed for a two-year period after the screening mammograms, and no patients were diagnosed with a breast malignancy.

The researchers randomly selected a subset of this group consisting of

4,855 patients (median age 54 years) broadly distributed across four ethnic/racial groups. The subset included 1,316 (27%) white, 1,261 (26%) Black, 1,351 (28%) Asian, and 927 (19%) Hispanic patients.



Example mammogram assigned a false-positive risk score of 1.0 in a 59-year-old Hispanic patient with heterogeneously dense breasts. Bilateral reconstructed two-dimensional (A, B) craniocaudal and (C, D) mediolateral oblique views are shown. The algorithm predicted cancer within 1 year, but this individual did not develop cancer or atypia within 2 years of the mammogram. Credit: Radiological Society of North America (RSNA)

A commercially available AI algorithm interpreted each exam in the subset of mammograms, generating both a case score (or certainty of malignancy) and a risk score (or one-year subsequent malignancy risk).

"Our goal was to evaluate whether an AI algorithm's performance was uniform across age, breast density types and different patient race/ethnicities," Dr. Nguyen said.

Given all mammograms in the study were negative for the presence of cancer, anything flagged as suspicious by the algorithm was considered a false positive result. False positive case scores were significantly more likely in Black and [older patients](#) (71-80 years) and less likely in Asian patients and younger patients (41-50 years) compared to white patients and women between the ages of 51 and 60.

"This study is important because it highlights that any AI software purchased by a health care institution may not perform equally across all patient ages, races/ethnicities and breast densities," Dr. Nguyen said.

"Moving forward, I think AI software upgrades should focus on ensuring demographic diversity."

Dr. Nguyen said health care institutions should understand the patient

population they serve before purchasing an AI algorithm for screening mammogram interpretation and ask vendors about their algorithm training.

"Having a baseline knowledge of your institution's demographics and asking the vendor about the ethnic and age diversity of their training data will help you understand the limitations you'll face in clinical practice," he said.

**More information:** Patient Characteristics Impact Performance of AI Algorithm in Interpreting Negative Screening Digital Breast Tomosynthesis Studies, *Radiology* (2024).

Provided by Radiological Society of North America

Citation: Age, race impact AI performance on digital mammograms, study finds (2024, May 21) retrieved 19 June 2024 from <https://medicalxpress.com/news/2024-05-age-impact-ai-digital-mammograms.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--