RSNA Press Release

AI Bias May Impair Radiologist Accuracy on Mammogram

Released: May 2, 2023

At A Glance

- Incorrect advice by an AI-based decision support system could impair the performance of radiologists when reading mammograms.
- When radiologists interpreted mammograms assisted by incorrect BI-RADS categories purportedly suggested by AI, their accuracy dropped sharply.
- The findings support the theory that use of AI assistance in mammography may make radiologists susceptible to automation bias.

OAK BROOK, Ill. — Incorrect advice by an AI-based decision support system could seriously impair the performance of radiologists at every level of expertise when reading mammograms, according to a new study published in *Radiology*, a journal of the Radiological Society of North America (RSNA).

download photo



Thomas Dratsch,
M.D., Ph.D.

Often touted as a "second set of eyes" for radiologists, AI-based mammographic support systems are one of the most promising applications for AI in radiology. As the technology expands, there are concerns that it may make radiologists susceptible to automation bias—the tendency of humans to favor suggestions from automated decision-making systems. Several studies have shown that the introduction of computer-aided detection into the mammography workflow could impair radiologist performance. However, no studies have looked at the influence of AI-based systems on the performance of accurate mammogram readings by radiologists.

Researchers from institutions in Germany and the Netherlands set out to determine how automation bias can affect radiologists at varying levels of experience when reading mammograms aided by an AI system.

In the prospective experiment, 27 radiologists read 50 mammograms. They then provided their Breast Imaging Reporting and Data System (BI-RADS) assessment assisted by an AI system. BI-RADS is a standard system used by radiologists to describe and categorize breast imaging findings. While BI-RADS categorization is not a diagnosis, it is crucial in helping doctors determine the next steps in care.

Researchers presented the mammograms in two randomized sets. The first was a training set of 10 in which the AI suggested the correct BI-RADS category. The second set contained incorrect BI-RADS categories, purportedly suggested by AI, in 12 of the 40 mammograms.

The results showed that the radiologists were significantly worse at assigning the correct BI-RADS scores for the cases in which the purported AI suggested an incorrect BI-RADS category. For example, inexperienced radiologists assigned the correct BI-RADS score in almost 80% of cases in which the AI suggested the correct BI-RADS category. When the purported AI suggested the wrong category, their accuracy fell to less than 20%. Experienced radiologists—those with more than 15 years of experience on average—saw their accuracy fall from 82% to 45.5% when the purported AI suggested the incorrect category.

"We anticipated that inaccurate AI predictions would influence the decisions made by radiologists in our study, particularly those with less experience," said study lead author Thomas Dratsch, M.D., Ph.D., from the Institute of Diagnostic and Interventional Radiology, at University Hospital Cologne in Cologne, Germany. "Nonetheless, it was surprising to find that even highly experienced radiologists were adversely impacted by the AI system's judgments, albeit to a lesser extent than their less seasoned counterparts."

The researchers said the results show why the effects of human-machine interaction must be carefully considered to ensure safe deployment and accurate diagnostic performance when combining human readers and AI.

"Given the repetitive and highly standardized nature of mammography screening, automation bias may become a concern when an AI system is integrated into the workflow," Dr. Dratsch said. "Our findings emphasize the need for implementing appropriate safeguards when incorporating AI into the radiological process to mitigate the negative consequences of automation bias."

Possible safeguards include presenting users with the confidence levels of the decision support system. In the case of an AI-based system, this could be done by showing the probability of each output. Another strategy involves teaching users about the reasoning process of the system. Ensuring that the users of a decision support system feel accountable for their own decisions can also help decrease automation bias, Dr. Dratsch said.

The researchers plan to use tools like eye-tracking technology to better understand the decision-making process of radiologists using AI.

"Moreover, we would like to explore the most effective methods of presenting AI output to radiologists in a way that encourages critical engagement while avoiding the pitfalls of automation bias," Dr. Dratsch said.

"Automation Bias in Mammography: The Impact of Artificial Intelligence BI-RADS Suggestions on Reader Performance." Collaborating with Dr. Dratsch were Xue Chen, M.D., Mohammad Rezazade Mehrizi, Ph.D., Roman Kloeckner, M.D., Aline Mähringer-Kunz, M.D., Michael Püsken, M.D., Bettina

Copyright ©2026 Radiological Society of North America (RSNA)

Baeßler, M.D., Stephanie Sauer, M.D., David Maintz, M.D., and Daniel Pinto dos Santos, M.D.

In 2023, *Radiology* is celebrating its 100th anniversary with 12 centennial issues, highlighting *Radiology*'s legacy of publishing exceptional and practical science to improve patient care.

Radiology is edited by Linda Moy, M.D., New York University, New York, N.Y., and owned and published by the Radiological Society of North America, Inc. (https://pubs.rsna.org/journal/radiology)

RSNA is an association of radiologists, radiation oncologists, medical physicists and related scientists promoting excellence in patient care and health care delivery through education, research, and technologic innovation. The Society is based in Oak Brook, Illinois. (RSNA.org)

For patient-friendly information on mammography, visit *RadiologyInfo.org*.

Images (JPG, TIF):

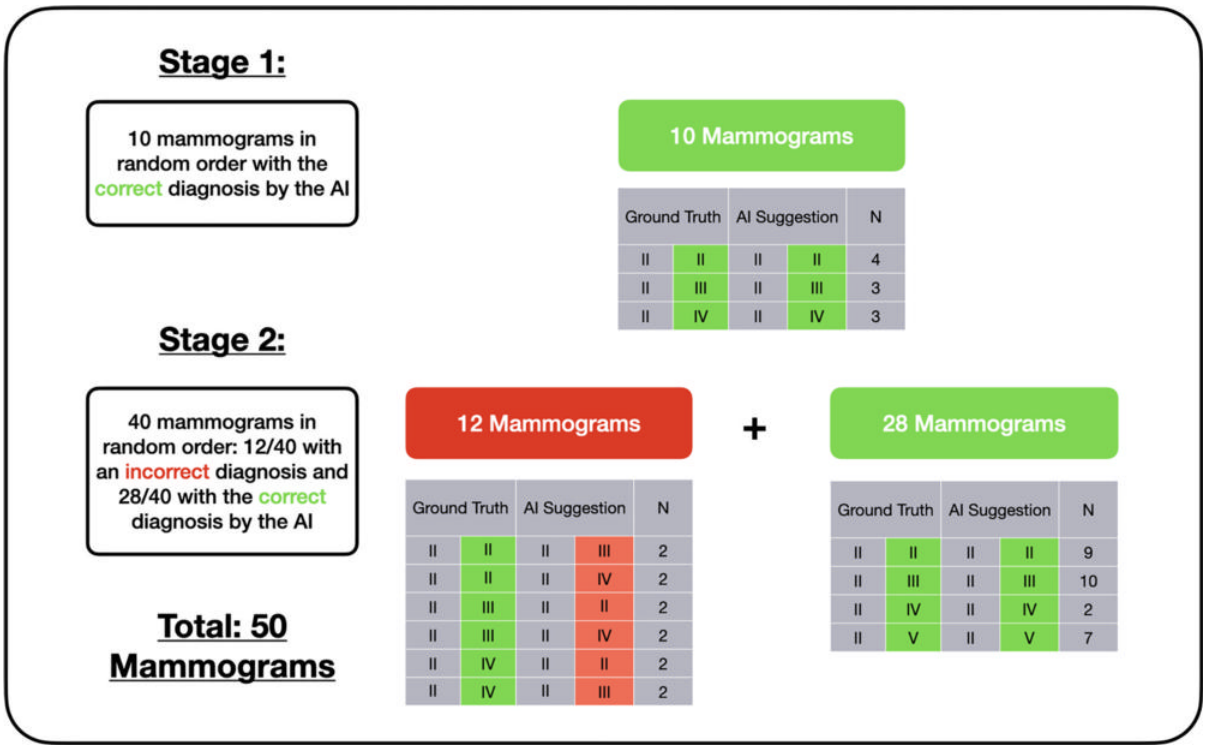


Figure 1. Design of the study and combinations of Breast Imaging Reporting and Data Systems categories used. AI = artificial intelligence.
High-res (TIF) version

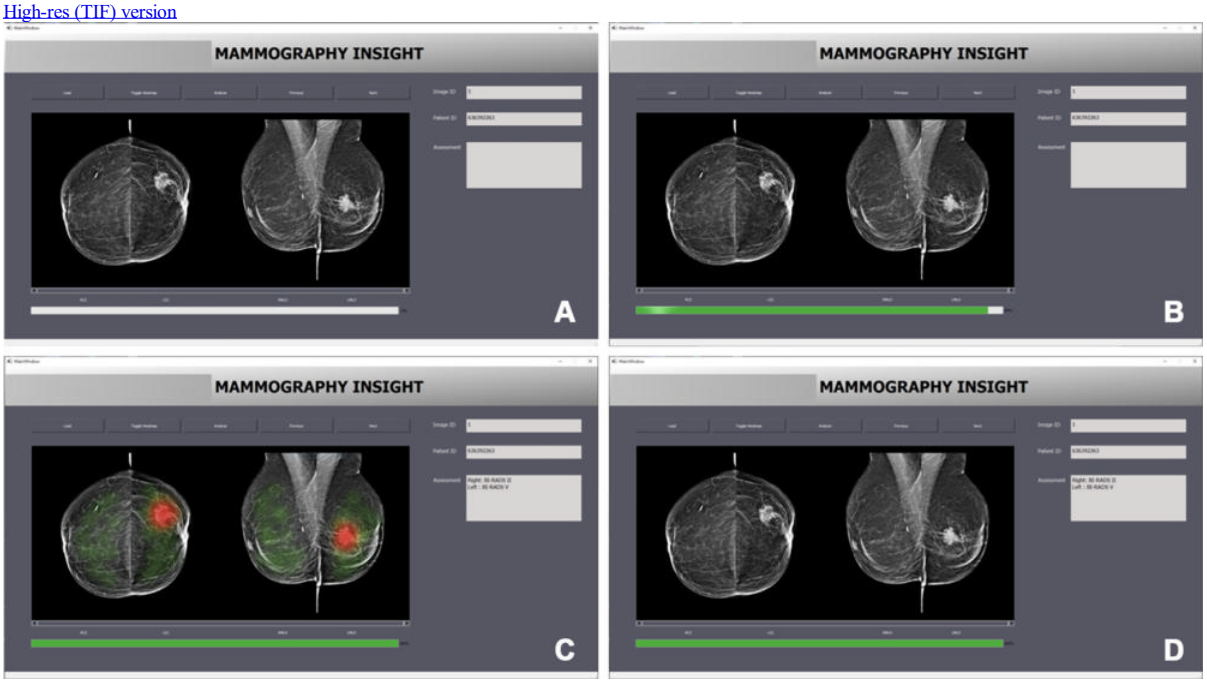


Figure 2. Interface of the artificial intelligence–based diagnostic system. Images show interface appearance when **(A)** loading a new mammogram, **(B)** performing the image evaluation, **(C)** displaying the results of the image evaluation via a heatmap, and **(D)** displaying the results of the image evaluation with the heatmap turned off.

Copyright ©2026 Radiological Society of North America (RSNA)

High-res (TIF) version

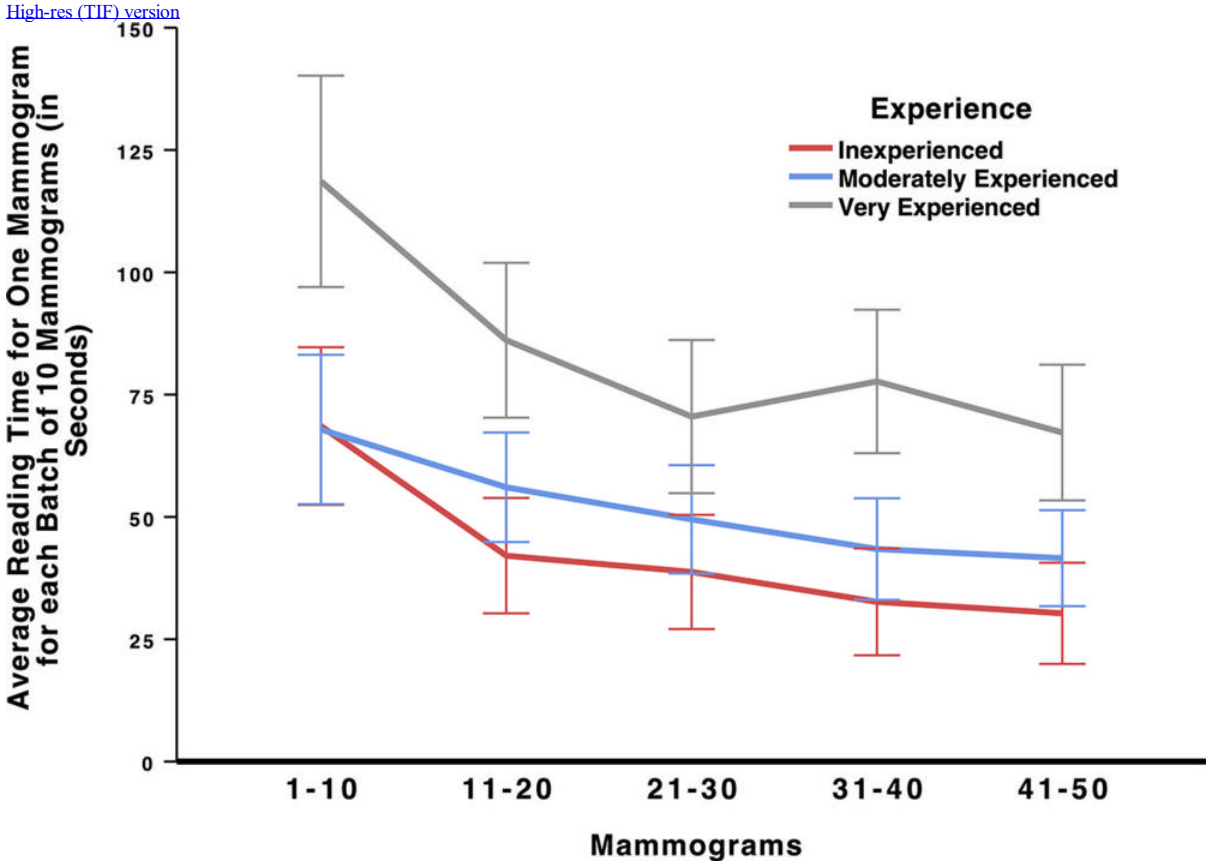


Figure 3. Line graph shows the average reading time (in seconds) for one mammogram for each batch of 10 mammograms for inexperienced (n = 11), moderately experienced (n = 11), and very experienced (n = 5) readers to evaluate mammogram images in the artificial intelligence–based diagnostic system over the course of the experiment. Average time to read mammograms decreased over time for both inexperienced and experienced readers. Error bars represent 95% CIs.

High-res (TIF) version

Copyright ©2026 Radiological Society of North America (RSNA)

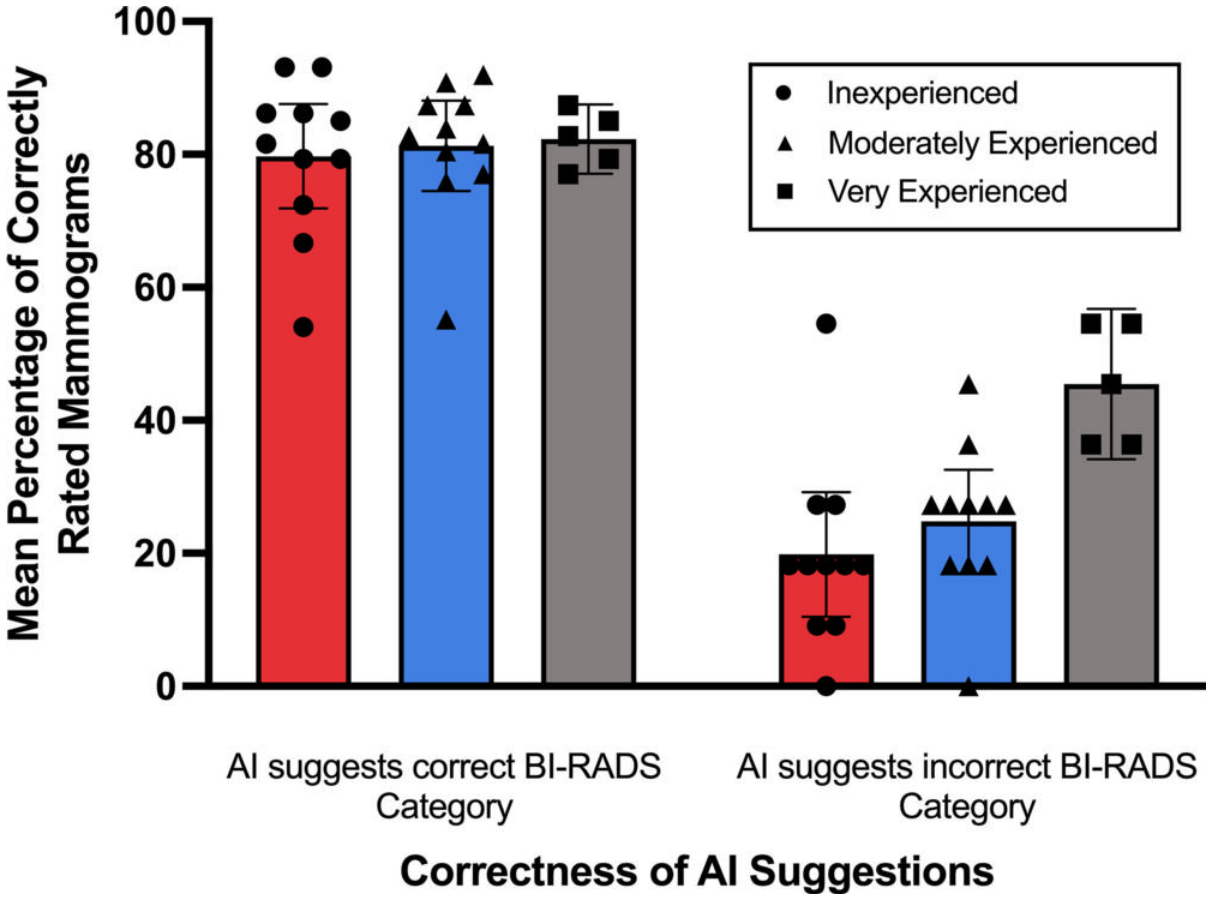


Figure 4. Bar graph shows the mean percentage of mammograms readers assigned the correct Breast Imaging Reporting and Data System (BI-RADS) score to stratified by correctness of artificial intelligence (AI) predictions and level of experience (inexperienced, $n = 11$; moderately experienced, $n = 11$; very experienced, $n = 5$). Error bars represent 95% CIs. Circles, triangles, and squares represent individual data points.
High-res (TIF) version

Copyright ©2026 Radiological Society of North America (RSNA)

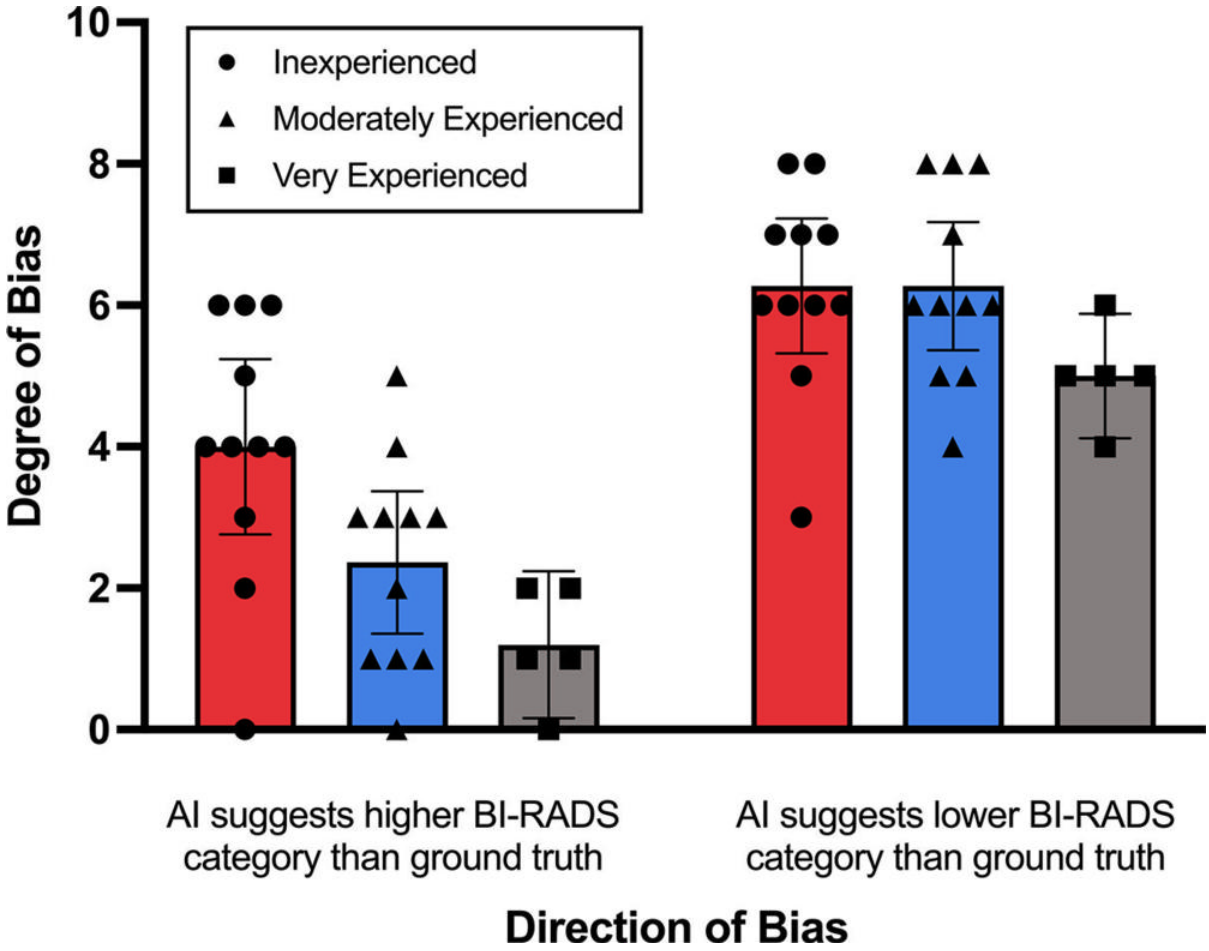


Figure 5. Bar graphs show the mean degree of bias stratified by correctness of artificial intelligence (AI) predictions and level of experience (inexperienced, *n* = 11, moderately experienced, *n* = 11, very experienced, *n* = 5). Error bars represent 95% CIs. Circles, triangles, and squares represent individual data points. BI-RADS = Breast Imaging Reporting and Data System.
High-res (TIF) version

Resources:

Editorial
Study abstract

Copyright ©2026 Radiological Society of North America (RSNA)