

Artificial Intelligence Used in Clinical Practice to Measure Breast Density

Released: October 16, 2018

At A Glance

- A new algorithm effectively measures breast density.
- Breast density can mask cancers on mammography and is an independent risk factor for the disease.
- The system has been in continuous operation at Massachusetts General Hospital since January and has processed approximately 16,000 images.

OAK BROOK, Ill. — An artificial intelligence (AI) algorithm measures breast density at the level of an experienced mammographer, according to a new study published in the journal *Radiology*. The researchers said the study, the result of a collaboration between breast imagers and AI experts, represents a groundbreaking implementation of AI into routine clinical practice.

[download full-size photo](#)



Constance D. Lehman, M.D., Ph.D.

Breast density can mask cancers on mammography and is an independent risk factor for the disease. The masking effect and cancer risk are significant enough that many states have laws mandating women be notified if they have mammographically dense breasts. Despite its importance, breast density assessment is an imperfect science, and research has shown much discrepancy among radiologists in making density determinations.

"We're dependent on human qualitative assessment of breast density, and that approach has significant flaws," said study lead author Constance D. Lehman, M.D., Ph.D., from Massachusetts General Hospital (MGH) in Boston. "We need a more accurate tool."

Dr. Lehman and colleagues collaborated with AI expert Regina Barzilay, Ph.D., professor of computer science and electrical engineering at the Massachusetts Institute of Technology in Cambridge, Mass., and her team to develop an algorithm that can automatically measure breast density. They used tens of thousands of high-quality digital mammograms from MGH to train and test the algorithm before implementing it in routine clinical practice. Eight radiologists then reviewed 10,763 mammograms that the model had determined were either dense or non-dense tissue.

The interpreting radiologist accepted the algorithm's assessment in 10,149 of the mammograms, or 94 percent. Dr. Lehman noted that the 94 percent agreement rate between the radiologists and the algorithm does not necessarily mean the machine was wrong in 6 percent of the cases. Reader variability could affect the disagreement because radiologists visually assess breast density, which is subjective and qualitative.

"We were thrilled with the results," Dr. Lehman said. "Now at Mass General, the deep learning algorithm processes all our screening mammograms and provides density, which is either accepted or rejected by the radiologists."

"The study results show that the algorithm worked remarkably well," Dr. Barzilay added. "But what's more important is that it is being used every day to measure breast density in mammograms at a major hospital."

The system has been in continuous operation at MGH since January and has processed approximately 16,000 images, according to Dr. Barzilay.

Dr. Lehman attributed the successful clinical implementation of the AI model to two components: the availability of high-quality, annotated data evaluated by expert radiologists, and the collaborative efforts of experienced, accomplished medical and computer science professionals.

"We have to have radiologists and other physicians who understand the pressing needs of our patients and can partner with computer scientists who are experts in AI," she said. "That is the collaboration that is going to move the field forward."

The algorithm has the potential to standardize and automate routine breast density assessment, the researchers said. On a broader scale, they see AI as central to the development of personalized breast cancer risk assessment for each woman who undergoes mammography. AI is uniquely suited to breast imaging because it can draw upon a large, mature database with advanced, structured reporting that links images with outcomes. This predictive power is a potential boon to all women, including groups who have been inadequately served by current prediction models. For instance, research has shown that current predictive models tend to underestimate risk in African-American women.

"We're teaching the machine to directly predict cancer risk even before the radiologist will see any cancer," Dr. Barzilay said. "The best time to control the disease is when it starts."

"With AI, we now have the ability to leverage vast amounts of information into more personalized, more targeted care for our patients," Dr. Lehman added. "In the case of breast cancer, we can better predict how likely a woman will have cancer in her future and improve the chances that it will be treated successfully."

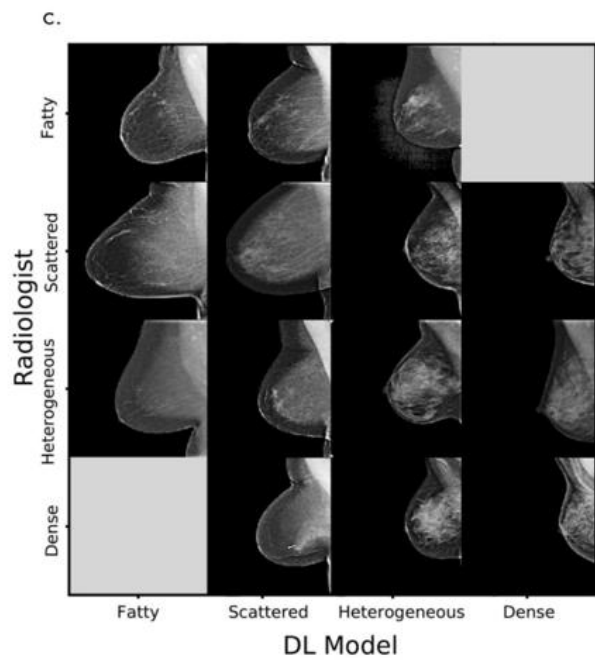
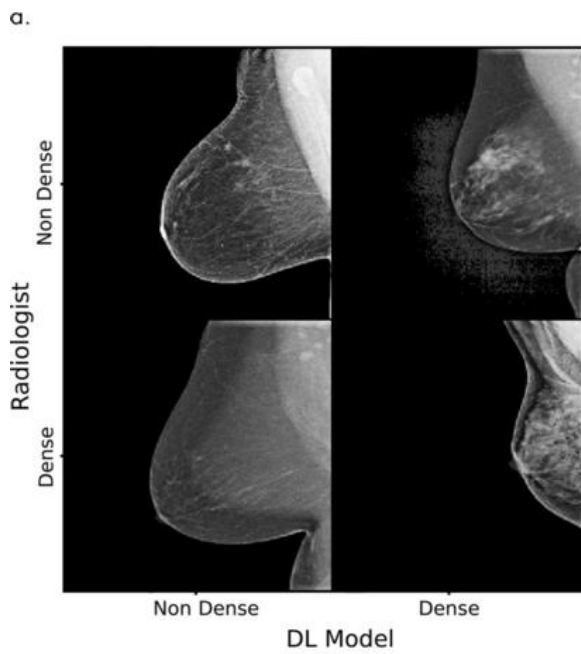
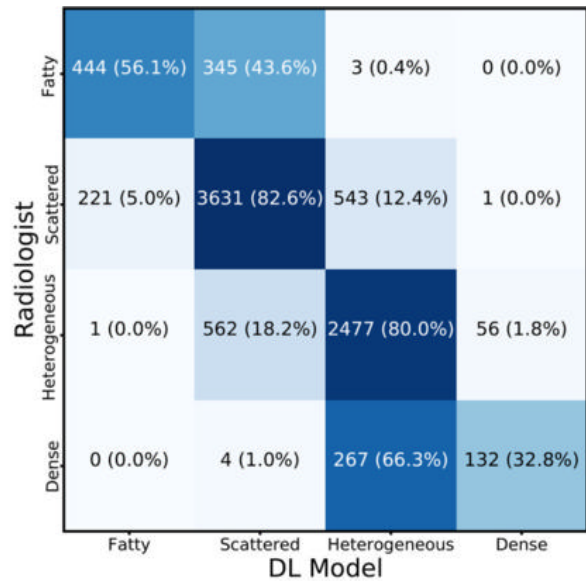
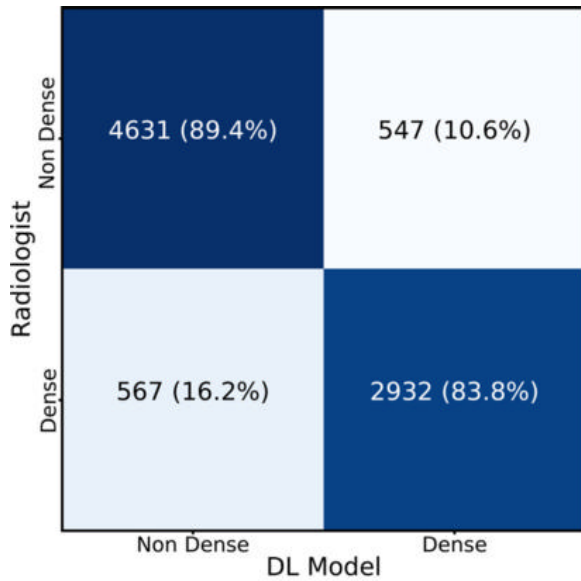
"Mammographic Breast Density Assessment using Deep Learning: Clinical Implementation." Collaborating with Drs. Lehman and Barzilay were Adam Yala, B.S., Tal Schuster, B.S., Brian Dontchos, M.D., Manisha Bahl, M.D., and Kyle Swanson, B.S.

Radiology is edited by David A. Bluemke, M.D., Ph.D., University of Wisconsin School of Medicine and Public Health, Madison, Wis., and owned and published by the Radiological Society of North America, Inc. (<http://radiology.rsna.org/>)

RSNA is an association of over 54,200 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, Ill. (RSNA.org)

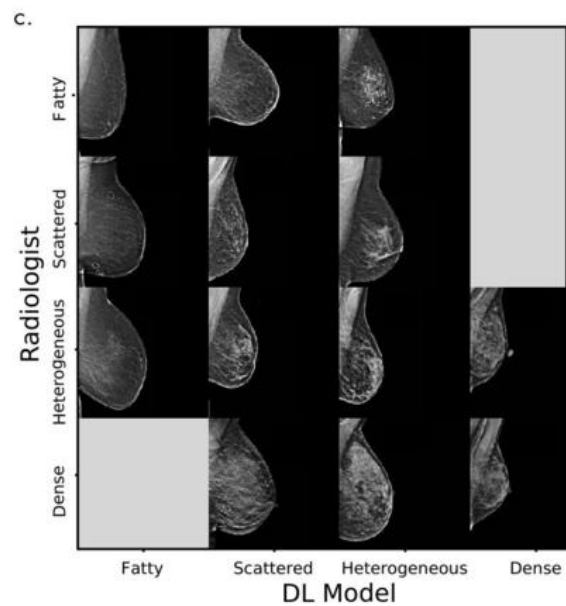
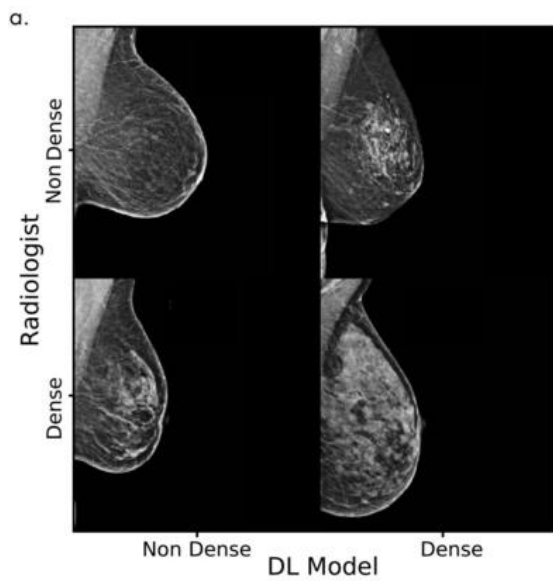
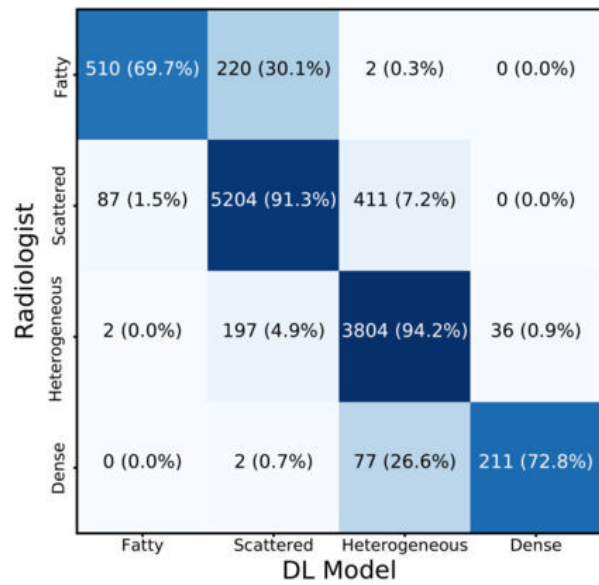
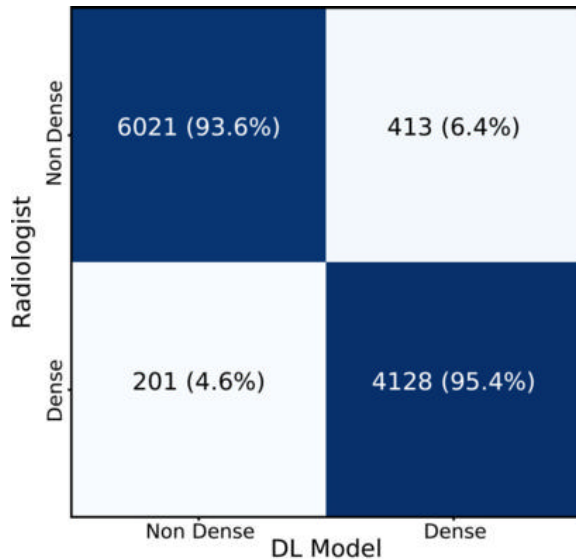
For patient-friendly information on breast imaging, visit RadiologyInfo.org.

Images (JPG, TIF):



b. d.
Figure 1. Test set assessment. Comparison of the original interpreting radiologist assessment with the deep learning (DL) model assessment for (a) binary and (c) four-way mammographic breast density classification. (b, d) Corresponding examples of mammograms with concordant and discordant assessments by the radiologist and with the DL model.

[High-res.\(TIF\) version](#)



b. d.
Figure 2. Clinical implementation assessment. Comparison of the original interpreting radiologist assessment with the deep learning (DL) model assessment for (a) binary and (c) four-way mammographic breast density classification. (b, d) Examples of mammograms with concordant and discordant assessments by the radiologist (b) and with the DL model (d).

[High-res \(TIF\) version](#)

Resources:

[Study abstract](#)