Editorial

eISSN 2005-8330
https://doi.org/10.3348/kjr.2021.0912
Korean J Radiol 2022;23(2):153-155

Korean Journal of Radiology
KJR

Research Highlight: Artificial Intelligence for Ruling Out Negative Examinations in Screening Breast MRI

Ji Hyun Youk[1], Eun-Kyung Kim[2]

[1]Department of Radiology, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Korea; [2]Department of Radiology, Yongin Severance Hospital, Yonsei University College of Medicine, Yongin, Korea

Take-home points

- Artificial intelligence (AI) application in triaging could optimize worklists after prioritizing examinations according to the complexity or likelihood of cancer diagnosis (i.e., soft triage) or identifying and ruling out negative or cancer-free examinations from the worklist (i.e., hard triage).
- A published AI model ruled out 39.7% of 3796 normal examinations maintaining 100% sensitivity for breast cancer in screening breast MRI for female aged 50–75 years with extremely dense breasts.
- Relative weights of the benefits and harms of dismissing examinations by AI systems should be taken into account.

Related KJR Quick Survey results are provided at the end of the article.

For breast cancer screening, mammography is the current standard modality for average-risk female, decreasing the mortality rate by detecting early-stage cancers. However, screening mammograms have some limitations, such as mammographically missed cancers. The sensitivity of screening mammography is 80%–85%, which declines to 50%–64% for extremely dense breasts [1,2]. Mammographically dense fibroglandular tissue that appears radiopaque as breast masses can reduce mammographic sensitivity by masking breast cancers. This has led to a growing interest in imaging modalities beyond mammography for supplemental screening, despite the ongoing debate on its benefits.

Breast MRI has excellent sensitivity for breast cancer detection and has been routinely offered as supplemental screening for high-risk populations only [3]. Nonetheless, it has demonstrated an increased detection rate in intermediate-risk and average-risk female [3,4]. MRI, not limited by mammographic breast density, may be valuable as a supplemental screening tool for average-risk female with dense breasts [5]. The Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial, a randomized controlled trial in the Netherlands, investigated the effect of supplemental MRI on the incidence of interval cancers in female with extremely dense breasts and showed its potential mortality benefit with an additional cancer detection rate of 16.5 per 1000 and a 50% reduction in interval cancer rate [5,6]. However, considering that extremely dense breasts account for approximately 10% of the screening population, radiologists' workload associated with the population-wide use of supplemental MRI for female with extremely dense breasts must be overwhelming [7].

Recently, deep-learning-based AI in breast imaging has been rapidly evolving, with promising solutions for diagnosis and decision support in breast cancer. In addition, the shortage of expert breast radiologists has sparked interest in the application of AI to reduce radiologists' workload, particularly in screening programs with the highest volume

Received: December 2, 2021 **Accepted:** December 17, 2021
Corresponding author: Eun-Kyung Kim, MD, Department of Radiology, Yongin Severance Hospital, Yonsei University College of Medicine, 363 Dongbaekjukjeon-daero, Giheung-gu, Yongin 16995, Korea.
• E-mail: ekkim@yuhs.ac

of examinations but a low prevalence of cancer (< 1%) [8]. The triaging system, which ranks patients based on urgency initially for emergency room settings, is now used in breast imaging to enhance efficiency while maintaining radiologists' performance. AI application in triaging could optimize worklists after prioritizing examinations according to the complexity or likelihood of cancer diagnosis (i.e., soft triage) or to identify and rule out negative or cancer-free examinations from the worklist (i.e., hard triage) [1,9-11].

Verburg et al. [12] recently published a study on AI application in automated triaging of 4581 supplemental breast MRI examinations in female with extremely dense breasts. They used a dataset from the first screening round of the DENSE trial for the AI algorithm to test the feasibility of dismissing normal MRI examinations (American College of Radiology Breast Imaging Reporting and Data System [BI-RADS] category 1) without any loss of sensitivity for detecting breast cancer. Their triaging model showed an average area under the receiver operating characteristic curve of 0.83 and ruled out 39.7% of 3796 normal MRI examinations, maintaining 100% sensitivity for breast cancer. A few other studies have reported the triage of screening mammograms using in-house deep-learning models or commercially available AI systems dedicated to cancer detection [13-16]. Optimal thresholds for triaging are supposed to be set regarding the trade-off between reducing the workload (i.e., dismissing normal, negative, or least suspicious examinations) and risking neglect cancers. Depending on the thresholds set in each study, the reduction in reading workload in screening mammography ranged from 17% to 91%. The AI models in these studies or typical commercial software AI tools are set to triage mammograms by choosing "cancer-free" thresholds. Therefore, it is unique that the model by Verburg et al. [12] was trained to triage "lesion-free" breast MRI examinations (BI-RADS category 1), not "cancer-free." It is inevitable that the model's performance is different from that of other AI systems for triaging mammograms. In addition, the threshold of the model by Verburg et al. [12] was set at a value corresponding to 100% sensitivity for cancer. Although Lång et al. [16] reported a 19.1% workload reduction by dismissing cancer-free mammographic examinations without missing any cancers, the remaining AI systems for triaging mammography reduced the workload at the expense of missing cancers. Although considered acceptable while maintaining the overall quality of outcomes, that is, with non-inferior performance to radiologists or screening programs, the relative weights of the benefits and harms of dismissing examinations by AI systems should be taken into account [14,16,17]. Moreover, the additional hemodynamic information to morphologic features provided by contrast-enhanced MRI examinations might result in a higher workload reduction (40%) of the model by Verburg et al. [12] than the mammographic AI system by Lång et al. [16] (19%), while the more severe background parenchymal enhancement was, the more normal MRI examination was triaged to radiologic review, which should be considered when applying the screening program for female aged < 50 years. In addition, medicolegal, regulatory, and ethical disputes remain challenging for pre-selecting examinations using a standalone AI system without human reading [9,16].

In the study by Verburg et al. [12], the model was trained on maximum intensity projection (MIP) images obtained by subtracting the pre-contrast image from the first post-contrast image of the dynamic contrast-enhanced MRI series. To overcome the limitations of full-protocol screening MRI, such as cost, patient tolerance, table time, and reading time, abbreviated breast MRI has been introduced as a screening examination, particularly in average-risk female [4,18]. Despite variations among institutions, the protocol for abbreviated breast MRI usually includes a pre-contrast image, the first post-contrast image, and their derived images (subtraction and MIP images) [3]. Therefore, the study results are compatible with and may enhance the efficiency of MRI screening programs using the abbreviated protocol. Notably, the study also provided a deep SHapley Additive exPlanations (SHAP) color map that overlaid MIP images with a color range from blue to red, indicating negative and positive SHAP values corresponding to a low or high probability of lesion presence, respectively. Dealing with the "black box" issue of AI, this is likely to increase the transparency and accountability of the AI model by allowing an intuitive and fast explanation of the AI decision [19].

Despite the promising work of the authors on AI triaging, limitations exist. The generalizability of the results is not guaranteed. All female included in the DENSE trial data study were from the Netherlands, and the performance of the AI triaging system would differ according to the demographic characteristics of specific screening populations, such as cancer prevalence. The model should be leveraged in various screening scenarios. As acknowledged by the authors, further application of AI

https://doi.org/10.3348/kjr.2021.0912

Korean Journal of Radiology
KJR

triaging is necessary for not only "lesion-free" but also "cancer-free" breast MRI.

Availability of Data and Material

Data sharing does not apply to this article as no datasets were generated or analyzed during the current study.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

ORCID iDs

Ji Hyun Youk
https://orcid.org/0000-0002-7787-780X
Eun-Kyung Kim
https://orcid.org/0000-0002-3368-5013

Funding Statement

None

REFERENCES

1. Bahl M. Artificial intelligence: a primer for breast imaging radiologists. *J Breast Imaging* 2020;2:304-314
2. Youk JH, Gweon HM, Son EJ, Eun NL, Kim JA. Fully automated measurements of volumetric breast density adapted for BIRADS 5th edition: a comparison with visual assessment. *Acta Radiol* 2021;62:1148-1154
3. Heacock L, Lewin AA, Toth HK, Moy L, Reig B. Abbreviated MR imaging for breast cancer. *Radiol Clin North Am* 2021;59:99-111
4. Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental breast MR imaging screening of women with average risk of breast cancer. *Radiology* 2017;283:361-370
5. Gao Y, Reig B, Heacock L, Bennett DL, Heller SL, Moy L. Magnetic resonance imaging in screening of breast cancer. *Radiol Clin North Am* 2021;59:85-98
6. Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *N Engl J Med* 2019;381:2091-2102
7. American College of Radiology. *ACR BI-RADS®-mammography ACR breast imaging reporting and data system*, 4th ed. Reston: American College of Radiology, 2003
8. Chan HP, Samala RK, Hadjiiski LM. CAD and AI for breast cancer-recent development and challenges. *Br J Radiol* 2020;93:20190580
9. Ou WC, Polat D, Dogan BE. Deep learning in breast radiology: current progress and future directions. *Eur Radiol* 2021;31:4872-4885
10. Kooi T. Cognitive biases and augmented intelligence in radiology. Lunit.io Web site. https://www.lunit.io/en/evidence/lunit-blog/cognitive-biases-and-augmented-intelligence-in-radiology. Accessed November 6, 2021
11. Yoon JH, Kim EK. Deep learning-based artificial intelligence for mammography. *Korean J Radiol* 2021;22:1225-1239
12. Verburg E, van Gils CH, van der Velden BHM, Bakker MF, Pijnappel RM, Veldhuis WB, et al. Deep learning for automated triaging of 4581 breast MRI examinations from the DENSE trial. *Radiology* 2022;302:29-36
13. Yala A, Schuster T, Miles R, Barzilay R, Lehman C. A deep learning model to triage screening mammograms: a simulation study. *Radiology* 2019;293:38-46
14. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Teuwen J, Broeders M, Gennaro G, et al. Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. *Eur Radiol* 2019;29:4825-4832
15. Kyono T, Gilbert FJ, van der Schaar M. Improving workflow efficiency for mammography using machine learning. *J Am Coll Radiol* 2020;17:56-63
16. Lång K, Dustler M, Dahlblom V, Åkesson A, Andersson I, Zackrisson S. Identifying normal mammograms in a large screening population using artificial intelligence. *Eur Radiol* 2021;31:1687-1692
17. Dembrower K, Wåhlin E, Liu Y, Salim M, Smith K, Lindholm P, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. *Lancet Digit Health* 2020;2:e468-e474
18. Ko ES, Morris EA. Abbreviated magnetic resonance imaging for breast cancer screening: concept, early results, and considerations. *Korean J Radiol* 2019;20:533-541
19. Morgan MB, Mates JL. Applications of artificial intelligence in breast imaging. *Radiol Clin North Am* 2021;59:139-148

 https://doi.org/10.3348/kjr.2021.0912

KJR Quick Survey: What Do You Think About Using AI to Screen Out "Normal/No Pathology" Examinations?

Survey period: October 14–31, 2021
Number of respondents: 85
Number of participating countries: 9

Do you agree with the general idea (not specifically for the breast MRI) of using AI to screen out "normal/no pathology" examinations and dismissing them?

- Yes: 27 (31.8%)
- Maybe: 45 (52.9%)
- No: 13 (15.3%)

If you answered No or Maybe to the previous question, what was the most important concern?

- Potential legal or ethical issues: 28 (48.3%)
- Not proven, reliable, or generalizable enough: 24 (41.4%)
- Threats to radiology jobs: 2 (3.4%)
- Others: 3 (5.2%), including multiple concerns (equally concerning) and neglecting major incidental findings
- No response: 1 (1.7%)

Who should take the largest responsibility for any harmful effects from significant missed lesions in the dismissed examinations?

- Doctor and/or hospital: 42 (49.4%)
- Regulatory bodies (such as FDA) that approved the use: 25 (29.4%)
- AI manufacturer: 17 (20%)
- No response: 1 (1.2%)