correspondence

Artificial Intelligence Risk Model (Mirai) Delivers Robust Generalization and Outperforms Tyrer-Cuzick Guidelines in Breast Cancer Screening

TO THE EDITOR:

Recently, Yala et al[1] reported on the validation of an artificial intelligence (AI)–aided breast cancer risk model called Mirai on the basis of traditional mammograms across multiple international cohorts. To the best of our knowledge, this is the first study suggesting that the AI-based breast cancer screening model developed on a large-scale data set across seven hospitals with at least five races can deliver impressively robust transferability. Such extensive cross-regional validation is challenging because of interethnic heterogeneity of breast cancer, device bias from different institutions, and completeness of patient clinical data.[2] To overcome these challenges, this study developed a risk factor predictor that could produce risk factors referred in the Tyrer-Cuzick (TC) guidelines and then combined these factors with mammographic images. This study also leveraged a conditional-adversarial training scheme to eliminate the bias caused by different mammography machines. As a result, Mirai maintained high performance (range, 0.75-0.84) and could accurately identify high-risk subsets across all test cohorts. Moreover, this AI model shows better performance than the widely used traditional TC model in improving early detection and reducing overtreatment of breast cancer. Thus, this study concluded that Mirai has the potential to replace current breast cancer risk assessment models recommended by clinical guidelines for magnetic resonance imaging screening.

Although this study demonstrated encouraging and promising findings, some issues should be noted. First, women with extremely dense breast tissue have double the risk of developing breast cancer as compared with the general screening population.[3] Meanwhile, the dense breast tissue limits the sensitivity of mammography,[4] causing a high number of interval cancers. Thus, we need to pay more attention to women with extremely dense breast tissue during breast cancer screening. However, the ability of the Mirai model in predicting breast cancer via mammography images presenting extremely dense breast tissue remains unclear. It would be clinically relevant to report the rate of extremely dense breast tissue in all data sets and conduct analysis to evaluate the performance of Mirai in predicting the risk of breast cancer in such patients. If Mirai has unsatisfactory performance, it should be leveraged with caution and the eligibility for supplemental magnetic resonance imaging or ultrasound should be determined to improve early detection and reduce interval cancers.[4,5] Second, risk prediction models are often evaluated by both calibration and discrimination, whereas this study used the area under the curve (AUC) and the concordance index (C-index) as discrimination metrics for evaluating the predictive performance of the Mirai model, which is not enough since calibration is not fully captured by AUC and C-index. The calibration reflects the consistency between the true probability and the model-predicted probability of cancer.[6] Calibration is a crucial property of a prediction model because it can assess the extent to which a model correctly estimates the absolute risk.[7] Models with poor calibration may overestimate the risk of cancer. Thus, both discrimination and calibration statistics are suggested to be used to evaluate the performance of the Mirai model. Third, this study used discriminative metrics AUC and C-index to compare the performance of the Mirai model and the TC model, whereas additional indicators such as integrated discrimination improvement and net reclassification index are helpful to evaluate the risk prediction improvement of the Mirai model compared with the TC model.[8] Finally, given the different results of Mirai between the populations screened annually and biennially, this model should provide appropriate risk thresholds to comply with local policies and procedures. In regions where screening is performed annually, Mirai should provide higher specificity to avoid supplemental screening, whereas in regions where screening is performed once every two or more years, Mirai should provide higher sensitivity to improve early detection.

Despite the abovementioned issues, we still appreciate Yala et al[1] for their outstanding work on the development and broadest validation of the Mirai model across race and ethnicity categories, which provided a more robust mammography-based breast cancer risk model. We look forward to further works to improve the validity and clinical utility of the Mirai model in real-world settings. The application of Mirai in breast cancer screening is still a long way from having the accuracy and generalizability required for its implementation into clinical practice.[9]

Zhe Jin, MD, Shuixing Zhang, PhD, Lu Zhang, PhD, Qiuying Chen, PhD, Shuyi Liu, PhD, and Bin Zhang, PhD

Department of Radiology, the First Affiliated Hospital of Jinan University, Guangdong, Guangzhou, China

EQUAL CONTRIBUTION

Z.J. and S.Z. contributed equally to this work.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at DOI https://doi.org/10.1200/JCO.21.02908.

REFERENCES

1. Yala A, Mikhael PG, Strand F, et al: Multi-institutional validation of a mammography-based breast cancer risk model. J Clin Oncol 40: 1732-1740, 2022
2. Yala A, Mikhael PG, Strand F, et al: Toward robust mammography-based models for breast cancer risk. Sci Transl Med 13:eaba4373, 2021
3. Price ER, Hargreaves J, Lipson JA, et al: The California Breast Density Information Group: A collaborative response to the issues of breast density, breast cancer risk, and breast density notification legislation. Radiology 269: 887-892, 2013
4. Bakker MF, de Lange SV, Pijnappel RM, et al: Supplemental MRI screening for women with extremely dense breast tissue. N Engl J Med 381: 2091-2102, 2019
5. Aripoli A, Fountain K, Winblad O, et al: Supplemental screening with automated breast ultrasound in women with dense breasts: Comparing notification methods and screening behaviors. AJR Am J Roentgenol 210: W22-W28, 2018
6. Liang W, Yang P, Huang R, et al: A combined nomogram model to preoperatively predict histologic grade in pancreatic neuroendocrine tumors. Clin Cancer Res 25:584-594, 2019
7. Alba AC, Agoritsas T, Walsh M, et al: Discrimination and calibration of clinical prediction models: Users' guides to the medical literature. JAMA 318: 1377-1384, 2017
8. Demler OV, Pencina MJ, Cook NR, et al: Asymptotic distribution of ΔAUC, NRIs, and IDI based on theory of U-statistics. Stat Med 36:3334-3360, 2017
9. Freeman K, Geppert J, Stinton C, et al: Use of artificial intelligence for image analysis in breast cancer screening programmes: Systematic review of test accuracy. BMJ 374:n1872, 2021

DOI: https://doi.org/10.1200/JCO.21.02908; Published at ascopubs.org/journal/jco on April 22, 2022.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Artificial Intelligence Risk Model (Mirai) Delivers Robust Generalization and Outperforms Tyrer-Cuzick Guidelines in Breast Cancer Screening

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/authors/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

No potential conflicts of interest were reported.