

Deep Learning Model Translates Imaging Biomarkers to Predict Future Breast Cancer Risk: Surpassing Traditional Methods of Risk Assessment

Monday, November 30, 8:30 – 9:30 AM CT | SSBR04-02 | Featured Paper

PURPOSE

A new deep learning (DL) algorithm was designed to predict a patient's risk of developing breast cancer at multiple time points using mammographic image biomarkers alone. The purpose of this study was to compare the predictive accuracy of the DL image-only model to that of Tyrer Cuzick version 8 (TC8), a traditional risk model that incorporates patient history and breast density to predict future breast cancer risk.

METHOD AND MATERIALS

The DL risk assessment model was developed using consecutive bilateral screening mammograms from 80818 patients between 2009 to 2016. Mammograms were randomly assigned to training, validation, or testing, resulting in 210819 examinations in 56831 patients, 25644 in 7021 patients, and 9290 in 3961 patients, respectively. No patients were excluded for model training and validation. For comparison of our 5-year DL model performance to TC8, women with a personal history of breast cancer, those who developed cancer within 3 months of the index screening mammogram, and those who lacked 5-year imaging follow-up were excluded. Cancer outcomes were obtained through linkage to a regional tumor registry. DL model vs TC8 model performance was compared using areas under the receiver operating characteristic curve (AUCs) with DeLong test ($p < 0.05$).

RESULTS

Mean patient age was 56.4 years (range 35 to 91). 6554/9290 (70.5%) were in post-menopausal patients and 2736/9290 (29.5%) were in pre-menopausal patients. 5170/9290 (55.7%) of patients had non-dense breasts (fatty/scattered fibroglandular) and 4116/9290 (44.3%) had dense breasts (heterogeneously dense/extremely dense). 7563/9290 (81.4%) were in white patients, 444/9290 (4.8%) in african american and 442/9290 (4.8%) in asian/pacific islander races. Race was unknown in 841/9290 (9.1%). The AUC of DL model was 0.71 (95% confidence interval [CI]: 0.676, 0.749) compared to 0.61 (95% CI: 0.572, 0.658) by TC8 model ($p < 0.001$).

CONCLUSION

Mammograms contain highly predictive biomarkers of future cancer risk, not identified by traditional risk models. A DL model using screening mammography alone can improve risk discriminatory accuracy compared to traditional modern risk models which rely on clinical history and mammographic breast density.

CLINICAL RELEVANCE/APPLICATION

Traditional risk models can be time-consuming to acquire and rely on inconsistent or missing data. A DL image-only risk model can provide increased access to more accurate, less costly risk assessment.