Machine learning can reliably predict malignancy of BI-RADS 4a and 4b breast lesions based on clinical and ultrasonographic features

Isabela Panzeri Carlotti Buzatto¹, Daniel Guimarães Tiezzi¹, Sarah Abud Recife², Ruth Morais Bonini³, Licerio Miguel², Liliane Silvestre¹, Nilton Onari³, Ana Luiza Peloso Araujo Faim³

¹Department of Obstetrics and Gynecology, Breast Disease Division, Ribeirão Preto Medical School, Universidade de São Paulo – Ribeirão Preto (SP), Brazil.
²Department of Gynecology & Obstetrics, Women’s Health Reference Center of Ribeirão Preto (MATER), Ribeirão Preto Medical School, Universidade de São Paulo – Ribeirão Preto (SP), Brazil.
³Department of Radiology, Hospital de Amor de Campo Grande – Campo Grande (MS), Brazil.

Objective: The objectives of this study were to establish the most reliable machine learning model to predict malignancy in BI-RADS 4a and 4b breast lesions and optimize the negative predictive value to minimize unnecessary biopsies.

Methodology: We included clinical and ultrasonographic attributes from 1,250 breast lesions from four institutions classified as BI-RADS 3, 4a, 4b, 4c, 5, and 6. We selected the most informative attributes to train the models in order to make inferences about the diagnosis of BI-RADS 4a and 4b lesions (validation dataset). Using the best parameters and hyper-parameters selected, we tested the performance of nine models and 1,530 ensemble models.

Results: The most informative attributes were shape, margin, orientation, and size of the lesions, the resistance index of the internal vessel, the age of the patient, and the presence of a palpable lump. The highest mean NPV was achieved with XGBoost (93.6%). The final performance of the best ensemble model was NPV=96.4%, sensitivity=81.5%, specificity=84.1%, PPV=46.8%, f1-score=59.5%, and the final accuracy=83.7%. Age was the most important attribute to predict malignancy. The use of the final model associated with the patient’s age would reduce by 51% the number of biopsies in women with BI-RADS 4a or 4b lesions.

Conclusion: Machine learning can predict malignancy in BI-RADS 4a and 4b breast lesions identified by ultrasonography, based on clinical and ultrasonographic features. Our final prediction model would be able to avoid 51% of the 4a and 4b breast biopsies, without missing any cancers.

Keywords: ultrasonography; machine learning; artificial intelligence; image-guided biopsy.