



Transparent Radiomics ML Model: Combining Human and Artificial Intelligence for Prediction of Therapy Outcomes

Shrey S. Sukhadia, MS, PhD, Dartmouth Health; Crisi Patel; Adrienne A. Workman, MD;
Roberta M. diFlorio-Alexander, MD; Marthony L. Robins, PhD

Introduction/Background

Predicting the response to neoadjuvant chemotherapy (NAC) in invasive breast carcinoma (IBC) is a key oncological challenge. Accurate prediction can help tailor individualized treatment, minimize chemotoxicity, and enhance therapeutic effectiveness. However, the efficacy of NAC varies significantly among patients, underscoring the importance of identifying reliable predictors of response. Advances in molecular biology, imaging, and artificial intelligence offer promising avenues for developing robust predictive models, which could revolutionize personalized treatment in breast cancer management. We identified the top nine latent radiomic features in the active tumor regions of pre-NAC MRI scans to predict the tumor's response to NAC using a transparent Decision Tree Classifier (DTC) model.

Methods/Intervention

We collected pre-NAC MRI scans for 75 IBC patients, for which the active tumor regions of interest (ROIs) underwent voxel-based segmentation using ITK-SNAP 4.0 (Fig. 1). The ROIs were fed to a custom built radiomic feature extraction pipeline that extracted 108 IBSI approved radiomic features from 7 feature classes and normalized using a standard scaler technique. Tumor response to NAC was extracted from the EHR with confirmed post NAC imaging with pathology report confirmation. An aggregated score was developed indicating tumor response to NAC. A DTC model was trained and tested using a 90:10 split of the sample-set using IMAGENE v3.2. A 3-fold cross validation was performed for training-set to control overfitting. The model was tested using the testing-set.

Results/Outcome

Our DTC model predicted the response to NAC at a remarkable AUC and R-square of 1.0 (each), at a p-value < 0.002. The radiomic features contributing the most to this prediction were shape, gray-level co-occurrence matrix, gray-level dependence matrix and gray-level size zone matrix. The Decision Tree showcasing the algorithm performed by DTC aided transparency of the model (Fig. 2).

Conclusion

We built a transparent Machine Learning model to predict NAC outcomes represented in EHR using the latent radiomic features extracted from MRIs pre-NAC. Our approach combines both human and artificial intelligence to identify novel radiomic biomarkers that predict therapy outcomes in breast cancer.

Statement of Impact

Our work offers a transparent Machine Learning model that predicts NAC outcomes using pre-NAC images in breast cancer.

Fig. 1

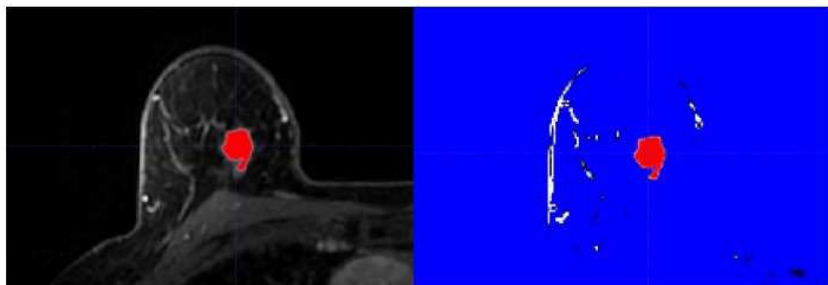


Fig. 1 Segmented Tumor ROIs on pre-NAC MRI scans using ITK-SNAP software

The diagram illustrates a hierarchical tree structure for signal decomposition. The root node is `original_glb1_SmallDependentLevelLevelDephases == 0.002`. It branches into `True` and `False`. The `True` branch leads to `original_shape_Dephases == 1.011`, which splits into `original_glb1_LargeDependentLevelLevelDephases == 0.19` and `original_shape_Dephases == 0.121`. The `False` branch leads to `original_glb1_SmallDependentLevelLevelDephases == 0.114`, which splits into `original_shape_Sparsity == 0.129` and `original_glb1_SmallDependentLevelLevelDephases == 2.191`. Each of these further splits into more nodes, eventually leading to 16 leaf nodes representing individual sub-signals with their parameters (glo, samples, value).

Keywords

Radiomics; Neoadjuvant Chemotherapy; Therapy Response; Machine Learning; MRI; Breast Cancer