Prediction of Oncotype DX recurrence score using deep multi-layer perceptrons in estrogen receptor positive, HER2 negative breast cancer

Aline Baltres[1], Zeina Al Masry[2], Ryad Zemouri[3], Severine Valmary-Degano[4], Laurent Arnould[5], Noureddine Zerhouni[2], Christine Devalland[1]

[1] Department of pathology, Nord Franche Comte hospital, Belfort, France.
[2] FEMTO-ST institute, Univ. Bourgogne Franche-Comté, CNRS, ENSMM, Besançon, France
[3] CEDRIC Laboratory of the Conservatoire National des Arts et Métiers (CNAM), HESAM Université, Paris, France.
[4] Department of pathology, university hospital, Grenoble, France.
[5] Department of pathology, Centre Georges-François Leclerc, Dijon, France

Abstract: Oncotype DX (ODX) is a multi-gene expression signature designed for estrogen receptor (ER)-positive and human epidermal growth factor receptor 2 (HER2)-negative breast cancer patients to predict the recurrence score (RS) and chemotherapy (CT) benefit. The aim of our study is to develop a prediction tool for the three RS's categories based on deep multi-layer perceptrons (DMLP) and using only the morphoimmunohistological variables. We performed a retrospective cohort of 320 patients who underwent ODX testing from three french hospitals. Clinico-pathological characteristics were recorded. We built a supervised machine learning classification model using Matlab software with 152 cases for the training and 168 cases for the testing. Three classifiers were used to learn the three risk categories of the ODX namely the low, intermediate and high risk. Experimental results provide the area under the curve (AUC) respectively for the three risk categories: 0.63 (95% confidence interval: [0.5446, 0.7154], $p < 0.001$), 0.59 (95% confidence interval: [0.5031, 0.6769], $p < 0.001$), 0.75 (95% confidence interval: [0.6184, 0.8816], $p < 0.001$). Concordance rate between actual RS and predicted RS was ranged from 53 to 56% for each class between DMLP and ODX. The concordance rate of low and intermediate combined risk group was 85%.

We developed a predictive machine learning model that could help to define patient's RS. Moreover, we integrated histopathological data and DMLP results to select tumor for ODX testing. Thus this process allows more relevant use of histopathological data and optimizes and enhances these information.

Keywords: Oncotype DX, Breast cancer, Deep multi-layer perceptrons, Prognostic factor, Histopathological feature.