DEEP LEARNING NEURAL NETWORK IMAGE ANALYSIS OF IMMUNOHISTOCHEMICAL PROTEIN EXPRESSION REVEALS A SIGNIFICANTLY REDUCED EXPRESSION OF BIGLYCAN IN BREAST CANCER

Ana Paula Thiesen[1], Bruna Mielczarski[2], Ricardo Francalacci Savaris[2]

[1]Universidade Federal do Rio Grande do Sul, Postgraduate Program in Medical Sciences – Porto Alegre (RS), Brazil.
[2]Universidade Federal do Rio Grande do Sul, Department of Gynecology and Obstetrics – Porto Alegre (RS), Brazil.

Objective: The aim of this study was to compare the protein expression of biglycan (BGN) in normal breast tissue and in breast cancer using deep learning and digital HScore techniques. **Methods:** In this case-control study, 24 formalin-fixed, paraffin-embedded tissues were obtained from pathological archives for analysis. Normal breast (n=9) and breast cancer (n=15) tissue sections were analyzed by immunohistochemistry using BGN monoclonal antibody (M01 – Abnova), clone 4E1-1G7 at dilution 1:300 at pH 6, and 3,3'-diaminobenzidine (DAB) as the chromogen. Photomicrographs of the slides were analyzed using the ImageJ software with "color deconvolution". After selecting the regions of interest (ROI), deconvoluted panels with DAB only were quantified using arbitrary DAB units. Another set, with higher magnification without ROI selection, was submitted to the inception V3 deep neural network image embedding recognition model. Next, supervised neural network analysis, using stratified 20-fold cross-validation, with 200 hidden layers, ReLu activation, and regularization at α=0.0001 were applied for SDLNN. The sample size was calculated for a minimum of seven cases and seven controls, having a power of 90%, an α error=5%, and a standard deviation of 20, to identify a decrease from the average of 40 DAB units (control) to 4 DAB units in cancer. Ethical approval was obtained from the Hospital de Clínicas de Porto Alegre Ethical Review Board (2019/0337). CAAE 15329119.9.0000.5327. **Results:** BGN expression (mean±SD) was 6.1±3.9 in breast cancer tissue, while in normal breast tissue, it was 39.6±21.9, using D-HScore (p=0.0017, student t-test, Welch corrected). SDLNN was able to correctly classify 110 out of 129 photomicrographs of the dataset using DAB panels only, with a classification accuracy of 85.3% (95%CI 78.1–90.3%) and the area under the curve of 94.3%. **Conclusion:** D-HScore and SDLNN revealed that BGN protein expression is reduced in breast cancer tissue, compared to normal tissue. The use of SDLNN seems to be a potential tool for image analysis in histological samples.

Keywords: Breast cancer. Biglycan. Immunohistochemistry. Deep learning. Digital HScore.