Introduction: Accurate and reproducible examination of predictive immunohistochemistry (IHC) assays is key for treatment selection in breast cancer. Local, conventional evaluation of IHC is subject to assay variability and interpretative subjectivity. Here, we applied an objective, automated platform for IHC quantification in a series of breast cancer biopsies and correlated its performance with traditional pathology assessment.

Methods: IHC slides for ER, HER2 and Ki-67 from 64 core biopsies from 60 patients diagnosed at a Chilean tertiary hospital were scanned using an Aperio® AT2 scanner (Leica). Whole tissue (WT) and tumor only (TO) were selected by an operator trained by a breast pathologist. These areas were analyzed using FDA-approved algorithms for nuclear and membrane positive pixel counting with capability for tumor cell detection. We compared agreement of WT and TO analysis to pathologist evaluation (PE) for percentage cell positivity (ER and Ki-67) and intensity scoring (HER2) using intraclass correlation (ICC) and kappa (κ) coefficients, respectively, as well as time required for analysis. All tests were two-sided (mean ± SEM).

Results: For all biomarkers, WT included an increased number of cells in the analysis, compared to strict TO selection (151550 ± 9773 vs 77567 ± 8050, P<0.0001). For
ER, WT and TO showed the highest concordance (ICC=0.924), while WT and PE had the lowest (ICC=0.837). For Ki-67, WT and TO showed the highest concordance (ICC=0.968), while WT and PE showed the lowest (ICC=0.84). When intensity scoring for HER2 was evaluated taking PE as standard, agreement was moderate and similar for WT and TO (κ =0.466 and 0.429, respectively). Automated TO evaluation required an increased amount of time per case, compared to WT and PE (124.6 ± 7.9 vs 4.7 ± 0.1 vs 6.9± 0.1 minutes, respectively, P<0.0001).

Conclusion: While objective IHC analysis shows promising results when compared to standard evaluation, issues with tissue recognition and duration of analysis need to be solved. These methodologies might find their best application in classifying borderline cases.