Enabling Molecular Subtyping in DCIS: Leveraging Foundation Models for Robust Biomarker Scoring and Patient Stratification

S. Doyle MS[†], M. Oerlemans MS[†], E. Marcus PhD, J. Brunekreef PhD, L. Mulder BS, Y. Liu MS, M. van Seijen PhD, I. Bouybayoune PhD, EJ. Sawyer PhD, AM Thompson MD, SE. Pinder PhD, Grand Challenge PRECISION Consortium, Prof Clara I. Sánchez PhD, J. Teuwen PhD*, Prof J. Wesseling PhD*, E. H. Lips PhD*

[†]Shared first authors

*Shared senior authors

Background

Ductal Carcinoma In Situ (DCIS) is a common breast pathology that may progress to ipsilateral invasive breast cancer. Accurate risk assessment is limited by poor understanding of progression and tumour heterogeneity. Biomarkers ER, HER2, and histological grade are associated with progression risk and guide inclusion in active surveillance trials for low-risk DCIS like the LORD study.

Research Question

Can deep learning models predict histological biomarkers (ER, HER2, grade) on whole-slide images (WSIs) for DCIS molecular subtyping and risk classification?

Methodology

From a Dutch nationwide cohort of 10,090 women diagnosed between 1989 and 2005, we selected 761 patients with primary, pure DCIS treated with breast-conserving surgery. FFPE blocks and clinico-pathological data were obtained through PALGA (Dutch Pathology Registry) and Dutch pathology labs. New HE-WSIs were produced, and graded by expert pathologists.
Biomarkers were binarized based on LORD criteria, defining low-risk DCIS as grade I/II, ER-positive, and HER2-negative. A deep learning pipeline with a foundation model encoder predicted biomarkers, using double nested k-fold cross-validation (k=5). Models were validated on the UK-based Sloane dataset (n=225).

Findings

On the Dutch dataset, models predicted ER, HER2, and grade with mean AUCs of 0.90, 0.84, and 0.86 and NPVs of 0.90 (±0.04), 0.83 (±0.01), and 0.81 (±0.03). External validation yielded AUCs of 0.80, 0.73, and 0.75 and NPVs of 0.80 (±0.02), 0.82 (±0.03), and 0.49 (±0.02). Risk stratification achieved AUCs of 0.73 (accuracy=0.73 with NPV=0.89) and 0.86 (accuracy=0.69 with NPV=0.84) on Dutch and external datasets, respectively.

Interpretation

Deep learning models accurately predicted biomarkers and stratified patients based on LORD low-risk criteria, demonstrating potential for integration in digitized pathology workflows, and furthering DCIS research.