Deep Learning for Predicting Invasive Recurrence of Ductal Carcinoma In Situ of the breast: Leveraging Histopathology Images and Clinical Features

S. Doyle MS, E. H. Lips PhD, E. Marcus 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*

*Shared senior authors

Background

Ductal Carcinoma In Situ (DCIS) can progress to ipsilateral invasive breast cancer (IBC) but over 75% of lesions would not progress and are therefore low-risk. Due to lack of reliable DCIS risk predictors, most patients receive IBC-like treatment, leading to overtreatment.

Research Question

Can deep learning models identify DCIS without invasive recurrence directly from histology whole-slide images (WSIs) and clinico-pathological data?

Methodology

From a Dutch nationwide cohort of 10,090 women diagnosed between 1989 and 2005, we selected 558 patients with primary, pure DCIS treated with breast-conserving surgery only. FFPE blocks and clinico-pathological data were obtained through PALGA (Dutch Pathology Registry) and the Dutch pathology labs. New HE-WSIs were produced, and graded by expert pathologists.
Using double nested k-fold cross-validation (k=5) on the Dutch dataset, we predicted invasive recurrence with Cox proportional hazards models using clinical data, and deep learning from WSIs alone or combined with clinical data (integrative). Models were tested on the UK-based Sloane dataset (n=94).

Results

Over 20 years, deep learning models using WSIs effectively stratified patients by invasive recurrence risk (negative predictive value (NPV)=0·79 (95% CI: 0·74–0·83); HR=4·48 (95% CI: 3·41–5·88, p<0·0001), AUC=0·75 (95% CI: 0·70–0·79)). Integrative models showed slightly higher hazard ratios ((NPV=0·79 (95% CI: 0·74–0·83); HR=4·48 (95% CI: 3·41–5·88, p<0·0001); AUC=0·75 (95% CI: 0·70–0·79). Clinical models achieved a less robust performance (NPV=0·64 (95% CI 0·59–0·69); HR=1·37 (95% CI 1·03–1·81, p=0·041); AUC=0·57 (95% CI 0·52–0·62)). External validation was unsuccessful due to the small dataset size and WSI quality.

Conclusion

Deep learning models using WSIs show promise for DCIS risk stratification. Integrating clinical data requires further study, and clinical factors alone cannot reliably predict recurrence. Generalization to larger cohorts could enable these models to reduce DCIS overtreatment.