

## **Title**

Development and Validation of a Transcriptomic Classifier for Predicting Invasive Breast Cancer Risk in Ductal Carcinoma In Situ

## **Authors**

Will Harley, Elinor J. Sawyer, Esther H. Lips, Maria Roman-Escorza, Tycho Bismeyer, Michael Sheinman, Ahmed A. Ahmed, Vandna Shah, Tapsi Kumar, Lodewyk F.A. Wessels, Jelle Wesseling

## **Background**

Ductal carcinoma in situ (DCIS) is a non-obligate precursor of invasive breast cancer (IBC), but less than 1% of DCIS cases progress to IBC annually. The inability to accurately predict progression has led to nearly all DCIS patients undergoing breast-conserving surgery (BCS) and radiotherapy. To combat overtreatment, we aimed to identify transcriptomic markers associated with progression risk and develop a prognostic tool to guide treatment decisions.

## **Research Question**

Can transcriptomic profiling of DCIS act as a prognostic tool to predict IBC risk and aid decision-making to de-escalate treatment?

## **Methods**

We compiled a population-based cohort of DCIS patients diagnosed between 1989 and 2005, through linkage of data between the Dutch Cancer Registry and Palga. From this cohort, we selected 186 primary pure DCIS samples treated by BCS without radiotherapy for RNA sequencing (RNA-seq). Clinical data was obtained by revision of pathology reports and FFPE blocks were obtained through Palga by the Dutch pathology labs. A machine learning classifier trained on this dataset identified 110 genes significantly associated with IBC risk. The classifier assigns a hazard ratio score based on the expression of these 110 genes, which determines the sample's assignment as low, intermediate or high risk. Performance was validated on independent RNA-seq data from the British Sloane Project (n=93).

## **Results**

Kaplan-Meier analysis shows significant differences in event-free survival between risk groups in both datasets censored at 5 and 10 years (e.g.,  $p = 1.77e^{-9}$  - Sloane dataset - 5 years). The negative predictive value in the low risk group is 98% and the ROC AUC is 0.78 at 5 years' follow-up in the Sloane set.

## **Conclusion**

The transcriptomic classifier effectively stratifies DCIS patients by IBC progression risk and demonstrates potential as a prognostic tool to reduce overtreatment of patients at low risk of progressing to IBC.